

Gene sets									
Patent: WO 0194629-A 6293 13-DEC-2001;									
Avalon Pharmaceuticals (US)									
Location/Qualifiers									
1. .3866									
/organism="Homo sapiens"									
/mol_type="unassigned DNA"									
/db_xref="taxon:9606"									
ORIGIN									
Query Match 100.0%; Score 3866; DB 6; Length 3866;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GGAAATCCCTTTT	TTTTTTT	TTTGGAGATGGAGTTT	CTCACTCTT	GTGGCCAGGCTGGAGTG	60		
Db	1	GGAAATCCCTTTT	TTTTTTT	TTTGGAGATGGAGTTT	CTCACTCTT	GTGGCCAGGCTGGAGTG	60		
QY	61	CAATGGCACAATCT	CAGCTT	ACTGCAACCT	CCGCCCTCCGGGTT	CAAGCGATTCTCCTGC	120		
Db	61	CAATGGCACAATCT	CAGCTT	ACTGCAACCT	CCGCCCTCCGGGTT	CAAGCGATTCTCCTGC	120		
QY	121	CTCAGCCTCTCAAG	TAGCTGGGATT	TACAGGCAT	GTGGCCACCACCCCT	GGCTAACTAATTT	180		
Db	121	CTCAGCCTCTCAAG	TAGCTGGGATT	TACAGGCAT	GTGGCCACCACCCCT	GGCTAACTAATTT	180		
QY	181	CTTTTCTATT	TAGTAGAGAT	GGGGTTT	CAACCAT	GTGGTCAGGCTGGTCTTGAACCTCCTG	240		
Db	181	CTTTTCTATT	TAGTAGAGAT	GGGGTTT	CAACCAT	GTGGTCAGGCTGGTCTTGAACCTCCTG	240		
QY	241	ACCTCAGGTGAT	CCACTT	GGCCTT	GGCCTCCCAAAGT	GCTAGGATTACAGCGGTGAAACTG	300		
Db	241	ACCTCAGGTGAT	CCACTT	GGCCTT	GGCCTCCCAAAGT	GCTAGGATTACAGCGGTGAAACTG	300		
QY	301	TGCCTGGCTGAT	CTTTT	TTTGTGTGGAT	TTTTT	GAAACAGGGTCTCCCTTGGTCGCC	360		
Db	301	TGCCTGGCTGAT	CTTTT	TTTGTGTGGAT	TTTTT	GAAACAGGGTCTCCCTTGGTCGCC	360		
QY	361	AGGCTGGAGTGC	AGTGGT	GCATCT	TGGCTCACTATAAACCT	CCACCTCGTGGTTCAAGT	420		
Db	361	AGGCTGGAGTGC	AGTGGT	GCATCT	TGGCTCACTATAAACCT	CCACCTCGTGGTTCAAGT	420		
QY	421	GATCCTCCCACT	TTTAGCCT	CTCAGTAGCT	GTGATTA	CAGGCGTGCAACCAACCCCGG	480		
Db	421	GATCCTCCCACT	TTTAGCCT	CTCAGTAGCT	GTGATTA	CAGGCGTGCAACCAACCCCGG	480		
QY	481	CTAATTTTGT	TATTTT	TATAGACACAGG	TTT	TACCATGTTTGGCCAGGCTGTTCTCAAA	540		
Db	481	CTAATTTTGT	TATTTT	TATAGACACAGG	TTT	TACCATGTTTGGCCAGGCTGTTCTCAAA	540		
QY	541	CTCCTGGACT	CAAGGGAT	CCGCTGCCT	CCACTTCCCAAAGT	TCCCGAGATTACAGGTGIG	600		
Db	541	CTCCTGGACT	CAAGGGAT	CCGCTGCCT	CCACTTCCCAAAGT	TCCCGAGATTACAGGTGIG	600		
QY	601	AGTCACCATG	CTGACCTT	ATAATTT	CTTAAGTCA	TTTTTCTGGTCCATTTCTTCCCTTAG	660		
Db	601	AGTCACCATG	CTGACCTT	ATAATTT	CTTAAGTCA	TTTTTCTGGTCCATTTCTTCCCTTAG	660		
QY	661	GGTCCTCACA	CAAAATCT	GCATTAGG	CGGTACAA	TAACTTCACTATGATTCACAAA	720		
Db	661	GGTCCTCACA	CAAAATCT	GCATTAGG	CGGTACAA	TAACTTCACTATGATTCACAAA	720		
QY	721	AGGAAGATGA	AGTGATTC	ATGATTT	TAGAAAGGGGAAGT	AGTAGAGCCCACTGCACACTCCT	780		
Db	721	AGGAAGATGA	AGTGATTC	ATGATTT	TAGAAAGGGGAAGT	AGTAGAGCCCACTGCACACTCCT	780		
QY	781	GGATGATGAT	CCTAAATCC	GATATAC	AGTAAATAGG	GATAGGAGGTAGAAATACAAA	840		
Db	781	GGATGATGAT	CCTAAATCC	GATATAC	AGTAAATAGG	GATAGGAGGTAGAAATACAAA	840		
QY	841	TTTGGTTTAA	ATTAATCT	ATAATCT	ATAAATCT	TTTGGATACATTTGTTGATGCTGA	900		
Db	841	TTTGGTTTAA	ATTAATCT	ATAATCT	ATAAATCT	TTTGGATACATTTGTTGATGCTGA	900		

QY	901	ATGTAAGACTGT	ACAGACTT	CTTAGAAAA	CAGTTT	GGGTCCATCT	TTTCAATTTCCCCAG	960		
Db	901	ATGTAAGACTGT	ACAGACTT	CTTAGAAAA	CAGTTT	GGGTCCATCT	TTTCAATTTCCCCAG	960		
QY	961	TGCAGTTTCT	GTAGAAAT	TGGAAT	CCGAGGAT	TTTAAGTGG	CAGAAATTGCAATTTGATT	1020		
Db	961	TGCAGTTTCT	GTAGAAAT	TGGAAT	CCGAGGAT	TTTAAGTGG	CAGAAATTGCAATTTGATT	1020		
QY	1021	CCATAATGA	CAAAAGT	GAGAGAC	ATTAATAA	TAAAGTTT	AAAAATGAAGACCTTACTGATG	1080		
Db	1021	CCATAATGA	CAAAAGT	GAGAGAC	ATTAATAA	TAAAGTTT	AAAAATGAAGACCTTACTGATG	1080		
QY	1081	AACTAAGCTT	GAATAAAA	TTTCTGCT	GATAC	TACAGATAA	CTCGGAACTGTTAAACCAA	1140		
Db	1081	AACTAAGCTT	GAATAAAA	TTTCTGCT	GATAC	TACAGATAA	CTCGGAACTGTTAAACCAA	1140		
QY	1141	TTATGATG	TGGCAAA	CAACC	CAGAGG	ACTGGTTG	AGTTTGTCTCAAACTAGAGAAA	1200		
Db	1141	TTATGATG	TGGCAAA	CAACC	CAGAGG	ACTGGTTG	AGTTTGTCTCAAACTAGAGAAA	1200		
QY	1201	ACAGTGTTC	CGCTAAGT	GTATGCT	TTTAAAT	TAAATTA	TGTTGTTGCTTACAGTCAAGCAA	1260		
Db	1201	ACAGTGTTC	CGCTAAGT	GTATGCT	TTTAAAT	TAAATTA	TGTTGTTGCTTACAGTCAAGCAA	1260		
QY	1261	TTGAAGCG	CTTCCCC	CAGATAA	TATGGCC	AAATGAG	AGTTTGTCTAGAAATTTCAAGTGA	1320		
Db	1261	TTGAAGCG	CTTCCCC	CAGATAA	TATGGCC	AAATGAG	AGTTTGTCTAGAAATTTCAAGTGA	1320		
QY	1321	GATTGCTGA	ATTAAA	AGCTATT	CAAGAG	CCAGATG	ATGCACTACTTTCAAATGG	1380		
Db	1321	GATTGCTGA	ATTAAA	AGCTATT	CAAGAG	CCAGATG	ATGCACTACTTTCAAATGG	1380		
QY	1381	CCAGAGCA	AACTGCA	AGAAAT	TTGCTT	TGTCATAT	ATCTTTTGCACAATTTGAACTGT	1440		
Db	1381	CCAGAGCA	AACTGCA	AGAAAT	TTGCTT	TGTCATAT	ATCTTTTGCACAATTTGAACTGT	1440		
QY	1441	CACAAGTA	ATGTCAA	AAAAAGT	AAACAA	CTTCTT	CAAAAAGCTGTAGAACGTGGAGCAG	1500		
Db	1441	CACAAGTA	ATGTCAA	AAAAAGT	AAACAA	CTTCTT	CAAAAAGCTGTAGAACGTGGAGCAG	1500		
QY	1501	TACCACTAG	AAATGCT	GGAAT	TGCCCT	CGGAAAT	TTTAAACCTCCAAAAAAGCAGCTGC	1560		
Db	1501	TACCACTAG	AAATGCT	GGAAT	TGCCCT	CGGAAAT	TTTAAACCTCCAAAAAAGCAGCTGC	1560		
QY	1561	TTTCAGAGG	AGGAAAA	AGAAAT	TTTAT	CAGCATCT	ACGGTATTAACCTCCCAAGAAATCAT	1620		
Db	1561	TTTCAGAGG	AGGAAAA	AGAAAT	TTTAT	CAGCATCT	ACGGTATTAACCTCCCAAGAAATCAT	1620		
QY	1621	TTTCCGGT	TCAC	TGGCAT	TTTAC	AGAAATAG	GAACAAACAGTTGTGATTCAGAGGACAGA	1680		
Db	1621	TTTCCGGT	TCAC	TGGCAT	TTTAC	AGAAATAG	GAACAAACAGTTGTGATTCAGAGGACAGA	1680		
QY	1681	CTACTAA	AGCCAGG	TTTAT	TATGG	AGAGACAT	GCCACCAAGATGCGCAATAGGTT	1740		
Db	1681	CTACTAA	AGCCAGG	TTTAT	TATGG	AGAGACAT	GCCACCAAGATGCGCAATAGGTT	1740		
QY	1741	ACCGGAAT	TCATTG	AGACAAA	ACTTAA	CAAACTP	AAACAGTCATGCCATTTGGAAGAGTCC	1800		
Db	1741	ACCGGAAT	TCATTG	AGACAAA	ACTTAA	CAAACTP	AAACAGTCATGCCATTTGGAAGAGTCC	1800		
QY	1801	CAGTTAAC	CTTCTA	ATATAG	CCAGAT	TTGTGAT	GTGAAGACAGATGATTCAGTTGTACCTT	1860		
Db	1801	CAGTTAAC	CTTCTA	ATATAG	CCAGAT	TTGTGAT	GTGAAGACAGATGATTCAGTTGTACCTT	1860		
QY	1861	GTTTTATG	AAAAGAC	AAACCT	CTAGAT	CAGAAAT	GCCGAGATTTGGTGTGCTGGATCTA	1920		
Db	1861	GTTTTATG	AAAAGAC	AAACCT	CTAGAT	CAGAAAT	GCCGAGATTTGGTGTGCTGGATCTA	1920		
QY	1921	AACCAAGT	GGAATG	ATTCCT	GTGAT	TAAAGT	CTTTTAAAGTCTGTTCAAAATAGTCATT	1980		
Db	1921	AACCAAGT	GGAATG	ATTCCT	GTGAT	TAAAGT	CTTTTAAAGTCTGTTCAAAATAGTCATT	1980		

QY 1981 TCAAGGAACCTCTGGTGCAGATGAAAAGAGTTCTGAACTTATTATTACTGATTCAATAA 2040
Db |||||
1981 TCAAGGAACCTCTGGTGCAGATGAAAAGAGTTCTGAACTTATTATTACTGATTCAATAA 2040
QY |||||
2041 CCCTGAAGAAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAAACCTAAAGAGTATC 2100
Db |||||
2041 CCCTGAAGAAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAAACCTAAAGAGTATC 2100
QY |||||
2101 AAGAACCGAGAGTTCCAGAGAGTAACCAAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db |||||
2101 AAGAACCGAGAGTTCCAGAGAGTAACCAAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
QY |||||
2161 GTATTAAACCGAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Db |||||
2161 GTATTAAACCGAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
QY |||||
2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
Db |||||
2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
QY |||||
2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTGAAGACACCAA 2340
Db |||||
2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTGAAGACACCAA 2340
QY |||||
2341 GCAGCAATACCTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGATGACT 2400
Db |||||
2341 GCAGCAATACCTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGATGACT 2400
QY |||||
2401 TTCCACCTGCTTGTGAGTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Db |||||
2401 TTCCACCTGCTTGTGAGTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
QY |||||
2461 AGCATCAAAATACCTGCCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
Db |||||
2461 AGCATCAAAATACCTGCCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
QY |||||
2521 ATGAATGCATTTTCGTTAAAGGAAGAAATTTATCCATATTAAAGCAGATAGGAAGTGGAG 2580
Db |||||
2521 ATGAATGCATTTTCGTTAAAGGAAGAAATTTATCCATATTAAAGCAGATAGGAAGTGGAG 2580
QY |||||
2581 GTTCAAGCAAGGTATTTCAAGTGTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
Db |||||
2581 GTTCAAGCAAGGTATTTCAAGTGTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
QY |||||
2641 TGAACCTTAGAAGAACGAGATAACCAACCTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db |||||
2641 TGAACCTTAGAAGAACGAGATAACCAACCTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
QY |||||
2701 TGAATAAACTACAACACAGTGTATAAGATCATCCGACTTTATGATTTATGAAATCAGG 2760
Db |||||
2701 TGAATAAACTACAACACAGTGTATAAGATCATCCGACTTTATGATTTATGAAATCAGG 2760
QY |||||
2761 ACCAGTACATCTACATGGTAAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAA 2820
Db |||||
2761 ACCAGTACATCTACATGGTAAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAA 2820
QY |||||
2821 AGAAAAATCCATTGATCCATGGGAACGCAAGATTACTGGAAAAATATGTTAGAGGCAG 2880
Db |||||
2821 AGAAAAATCCATTGATCCATGGGAACGCAAGATTACTGGAAAAATATGTTAGAGGCAG 2880
QY |||||
2881 TTCAACAATCCATCAACATGGCATTTGTTACAGTGTATCTTAAACCCAGCTAACTTTCTGA 2940
Db |||||
2881 TTCAACAATCCATCAACATGGCATTTGTTACAGTGTATCTTAAACCCAGCTAACTTTCTGA 2940
QY |||||
2941 TAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTCAAACCAATGCAACCCAGATA 3000
Db |||||
2941 TAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTCAAACCAATGCAACCCAGATA 3000
QY |||||
3001 CAACAAGTGTGTTAAAGATTCTCAGGTTGGCAGAGTAAATATATGACCAGAGCAAC 3060
Db |||||
3001 CAACAAGTGTGTTAAAGATTCTCAGGTTGGCAGAGTAAATATATGACCAGAGCAAC 3060
QY |||||
3061 TCAAGATATGTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATAAGCCCCAAAA 3120

Db |||||
3061 TCAAGATATGTCTTCTCCAGAGAGAAATGGAAATCTAAGTCAAGATAAGCCCCAAAA 3120
QY |||||
3121 GTGATGTTTGGTCTTAGGATGTATTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db |||||
3121 GTGATGTTTGGTCTTAGGATGTATTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
QY |||||
3181 AGCAGATAATTAATCAGATTTCTAAATTACATGCCATAAATTGATCCTAATCATGAATTG 3240
Db |||||
3181 AGCAGATAATTAATCAGATTTCTAAATTACATGCCATAAATTGATCCTAATCATGAATTG 3240
QY |||||
3241 AATTTCCCGATATCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db |||||
3241 AATTTCCCGATATCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
QY |||||
3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAA 3360
Db |||||
3301 ACCCAAAACAGAGATATCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAA 3360
QY |||||
3361 CTCATCCAGTTAACCAAAATGGCCAAAGGAACCACTGAAGAAATGAATAATGTTCTGGGCC 3420
Db |||||
3361 CTCATCCAGTTAACCAAAATGGCCAAAGGAACCACTGAAGAAATGAATAATGTTCTGGGCC 3420
QY |||||
3421 AACTTGTGGTCTGAATTTCTTAACCTCAATTTGAAAGCTGCTAAACCTTTTATATGAAC 3480
Db |||||
3421 AACTTGTGGTCTGAATTTCTTAACCTCAATTTGAAAGCTGCTAAACCTTTTATATGAAC 3480
QY |||||
3481 ACTATAGTGGTGGTGAAGTCAATAATTTCTCATCTCCAAGACTTTTGAAAAAAAAGGG 3540
Db |||||
3481 ACTATAGTGGTGGTGAAGTCAATAATTTCTCATCTCCAAGACTTTTGAAAAAAAAGGG 3540
QY |||||
3541 GAAAAAAATGATTTGCACTTATTCGTAATGTCAGATAGGAGGTATAAAAATATATTTGGACT 3600
Db |||||
3541 GAAAAAAATGATTTGCACTTATTCGTAATGTCAGATAGGAGGTATAAAAATATATTTGGACT 3600
QY |||||
3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACTCACTCTGAAGTGT 3660
Db |||||
3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACTCACTCTGAAGTGT 3660
QY |||||
3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAACTGTAAAAATAGCAACCAC 3720
Db |||||
3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAACTGTAAAAATAGCAACCAC 3720
QY |||||
3721 TTATGGCACTGTATATATTTGAGACTTGTGTTTCTCTGTTTATGCTCTTGTGTATCTAC 3780
Db |||||
3721 TTATGGCACTGTATATATTTGAGACTTGTGTTTCTCTGTTTATGCTCTTGTGTATCTAC 3780
QY |||||
3781 TTGACATCATTTTACTCTTTGGAATAGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTG 3840
Db |||||
3781 TTGACATCATTTTACTCTTTGGAATAGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTG 3840
QY |||||
3841 TAAATAAAGTTTTGTGGCTAAAAATGA 3866
Db |||||
3841 TAAATAAAGTTTTGTGGCTAAAAATGA 3866

RESULT 2

HUMTTK Human kinase (TTK) mRNA, complete cds. mRNA linear PRI 14-JAN-1995
LOCUS M86699.1 GI:340010
DEFINITION binding protein; kinase; regulatory protein.
ACCESSION M86699
VERSION 1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3866)
AUTHORS Mills,G.B., Schmandt,R., McGill,M., Amendola,A., Hill,M.,
Jacobs,K., May,C., Rodricks,A.M., Campbell,S. and Hogg,D.
TITLE Expression of TTK, a novel human protein kinase, is associated with
cell proliferation
JOURNAL J. Biol. Chem. 267 (22), 16000-16006 (1992)

QY 1381 CCAGAGCAAACGCAAGAAATTTGCTTTGTTTCATATATCTTTTGACAAATTTGAACTGT 1440
Db 1381 CCAGAGCAAACGCAAGAAATTTGCTTTGTTTCATATATCTTTTGACAAATTTGAACTGT 1440
QY 1441 CACAAGGTAATGTCAAAAAAAGTAACAACCTTCTTCAAAAAAGCTGTAGAACGTCGAGCAG 1500
Db 1441 CACAAGGTAATGTCAAAAAAAGTAACAACCTTCTTCAAAAAAGCTGTAGAACGTCGAGCAG 1500
QY 1501 TACCACCTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGC 1560
Db 1501 TACCACCTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGC 1560
QY 1561 TTTTCAGAGGAGGAAAGAAATTTTATCAGCATCTACGGTATTAACTGCCCAAGATCAT 1620
Db 1561 TTTTCAGAGGAGGAAAGAAATTTTATCAGCATCTACGGTATTAACTGCCCAAGATCAT 1620
QY 1621 TTTCCGGTTCACTTCGGCATTTTACAGATAGGAACAACAGTTGTGATTTCCAGAGGACAGA 1680
Db 1621 TTTCCGGTTCACTTCGGCATTTTACAGATAGGAACAACAGTTGTGATTTCCAGAGGACAGA 1680
QY 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTT 1740
QY 1741 ACCGGAATTCATTGAGACAAAACCTTAACAACTAAACAGTCATGCCCATTTGGAAAGAGTCC 1800
Db 1741 ACCGGAATTCATTGAGACAAAACCTTAACAACTAAACAGTCATGCCCATTTGGAAAGAGTCC 1800
QY 1801 CAGTTAACCTTCTAAATAGCCAGATTTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAACCTTCTAAATAGCCAGATTTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
QY 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCCTGGATCTA 1920
QY 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGGTGTGATGAAAAGAGTCTCTGAACCTTATTTACTGATTCATAA 2040
Db 1981 TCAAGGAACCTCTGGTGTGATGAAAAGAGTCTCTGAACCTTATTTACTGATTCATAA 2040
QY 2041 CCTGGAAGATTAACCGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
Db 2041 CCTGGAAGATTAACCGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
QY 2101 AAGAACAGAGGTTCCAGAGAGTAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAT 2160
Db 2101 AAGAACAGAGGTTCCAGAGAGTAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAT 2160
QY 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
QY 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTTGTAAGACACCAA 2340
Db 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTTGTAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTGAACCTCCAGTGTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTGAACCTCCAGTGTAAAGAAATGACT 2400
QY 2401 TTCACCTGCTTGTGAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCACCTGCTTGTGAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460

QY 2461 AGCATCAAATACCTTGCCACTCCACCTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAATACCTTGCCACTCCACCTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
QY 2521 ATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGATTTTCAGGTGTTAAATGAAAAGAAAACAGATATATGCTATAAAAATATG 2640
Db 2581 GTTCAAGCAAGGATTTTCAGGTGTTAAATGAAAAGAAAACAGATATATGCTATAAAAATATG 2640
QY 2641 TGAACCTTAGAAGAACAGATTAACCAAACCTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAACAGATTAACCAAACCTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAACAACAGATGAAGATCATCCGACTTTATGATTAATAAGATCAGCG 2760
Db 2701 TGAATAAACTACAACAACAGATGAAGATCATCCGACTTTATGATTAATAAGATCAGCG 2760
QY 2761 ACCAGTACATCTACATGTAATGGAGTGTGAAAATATTGATCTTAAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGTAATGGAGTGTGAAAATATTGATCTTAAATAGTTGGCTTAAAA 2820
QY 2821 AAAAAAAATCCATTCATCCATGGGAACGCAAGAGTTACTGAAAAAATATGTTAGAGGCAG 2880
Db 2821 AAAAAAAATCCATTCATCCATGGGAACGCAAGAGTTACTGAAAAAATATGTTAGAGGCAG 2880
QY 2881 TTCCACAAATCCATCAACATGGCATTTGTTCCAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Db 2881 TTCCACAAATCCATCAACATGGCATTTGTTCCAGTGATCTTAAACCAGCTAACTTTCTGA 2940
QY 2941 TAGTTGATGGAATGCTTAAAGCTAAATTTGATTTTGGGATTTGCAAAACCAATGCAACCAGATA 3000
Db 2941 TAGTTGATGGAATGCTTAAAGCTAAATTTGATTTTGGGATTTGCAAAACCAATGCAACCAGATA 3000
QY 3001 CAACAAGTGTGTTTAAAGATTTCTCAGGTTGGCACAGTTAATTATATGSCACCAGAAAGCAA 3060
Db 3001 CAACAAGTGTGTTTAAAGATTTCTCAGGTTGGCACAGTTAATTATATGSCACCAGAAAGCAA 3060
QY 3061 TCAAGATATGTCCTTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCCCAAAA 3120
Db 3061 TCAAGATATGTCCTTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCCCAAAA 3120
QY 3121 GTGATGTTTGGTCCTTAGGATGATTTTGTACTATATGACTTACGGGAAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCCTTAGGATGATTTTGTACTATATGACTTACGGGAAAACACCATTTTC 3180
QY 3181 AGCAGATAATTAATCAGATTTCTAAATTAATGATCCCATTAATGATCCTTAATCATGAAATTG 3240
Db 3181 AGCAGATAATTAATCAGATTTCTAAATTAATGATCCCATTAATGATCCTTAATCATGAAATTG 3240
QY 3241 AATTTCCTCGATATTCAGAGAAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCTCGATATTCAGAGAAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
QY 3301 ACCCAAAAACAGAGGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAA 3360
Db 3301 ACCCAAAAACAGAGGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAA 3360
QY 3361 CTCATCCAGTTAACCAAAATGGCCAAAGGAAACCACTGAAGAAATGAAAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAACCAAAATGGCCAAAGGAAACCACTGAAGAAATGAAAATATGTTCTGGGCC 3420
QY 3421 AACTTGTGTTGTTCTGAATTCCTTAACCTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
Db 3421 AACTTGTGTTGTTCTGAATTCCTTAACCTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
QY 3481 ACTATAGTGGTGAAGTCAATAATCTTCTCATCTCCAGACTTTTGAAGAAAAAAGGG 3540
Db 3481 ACTATAGTGGTGAAGTCAATAATCTTCTCATCTCCAGACTTTTGAAGAAAAAAGGG 3540
QY 3541 GAAAAAATGATTTGTCAGTTATTTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600

Db 3541 GAAAAAATGATTGAGTTTATTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600
QY 3601 GTTATACCTTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db 3601 GTTATACCTTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAAACTGTAAAAATAGCAACCAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAAACTGTAAAAATAGCAACCAC 3720
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTCTCTCTGTTTATGCTCTGTGTAACTCTAC 3780
Db 3721 TTATGGCACTGTATATATTGTAGACTTGTCTCTCTGTTTATGCTCTGTGTAACTCTAC 3780
QY 3781 TTGACATCATTTTACTCTTGGATAGTGGGTGGATAGCAAGTATATCTTAAAAAACTTTG 3840
Db 3781 TTGACATCATTTTACTCTTGGATAGTGGGTGGATAGCAAGTATATCTTAAAAAACTTTG 3840
QY 3841 TAAATAAAGTTTGTGGCTAAATGA 3866
Db 3841 TAAATAAAGTTTGTGGCTAAATGA 3866

RESULT 3
BC032858 2980 bp mRNA linear PRI 12-NOV-2003
LOCUS Homo sapiens TTK protein kinase, mRNA (cdna clone MGC:33286
DEFINITION IMAGE:5270012), complete cds.
ACCESSION BC032858
VERSION BC032858.2 GI:34190560
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2980)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
human and mouse cdna sequences
Generation and initial analysis of more than 15,000 full-length
human and mouse cdna sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
REFERENCE 2 (bases 1 to 2980)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT On Aug 25, 2003 this sequence version replaced gi:23271248.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cdna Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 47 Row: p Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1. .2980
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:33286 IMAGE:5270012"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"
1. .2980
/gene="TTK"
/note="synonyms: MPS1L1, PYT, ESK"
/db_xref="LocusID:7272"
/db_xref="MIM:604092"
77. .2650
/codon_start=1
/product="TTK protein kinase"
/protein_id="AAH32858.1"
/db_xref="GI:23271249"
/db_xref="LocusID:7272"
/translation="MESEDLGRELITDSIMNKVRDIKNKPKNEDLTDELSLNKISAD
TTDNGTGNQIMMANNPEDWLSLLKLEKNSVPLSDALLNKLIGRYSQAIEALPPDK
YQNESFARIQVFAELKAIQEPDDARDYFQMARANCKKFAFVHISFAQFELSQCNVK
KSKQLLQKAVERGAVPELEMLEIALNLQKQLLSEEEKNLASTVLTAEQSFSGS
LGHQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYRNSLRQTNKTQSCPFGRVPEV
NLLNSPCDVKTDDSVVPCFMKQTSRSECRDLVVPKSPGSGNDSCELRNLKSVQNSH
FKEPLVSDKESSELIITDSITLKNKTESLLAKLEETKEYQEPVEPESNQKQWSKRRK
SECINQNPAASSNHQIPELARKVNTQKHTTTEQPVFVSQSPPISTSKWFDPKSI
CKTPSSNTLDDYMSCFRTPVKNDFPACQLSTPYQPCFQQQHQLATPLQNLQV
LASSANECISVKGRISYILKQIGSGSSKVFQVLEKQIYAIKYVNLEEDNQTLN
SYNEIAYLNKLQHSDKIIRLYDEITDQIYVMMECIGNIDLSWLKRSIDPWER
KSYWKMLEAVHTIHQHIVHSDLEKPFANFLIVDMLKIDFGIANQMOPDTTSVVKDS
QVGTVMPPPEAKDMSSRENGSKSKISPKSDVWSLGLIYMTYGTPTFOQIINQ
ISKLHAIIDPNHEIEFPDIPEKDLQVLKCLKRPKQKRIPELLAHPYVQIOTHPV
NQMAKGTTEEMKVLQVLGLNSPILKAATLYEHYSGESHSSSSSKTFEKKRGK
K"

FEATURES
source
1. .2980
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:33286 IMAGE:5270012"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"
1. .2980
/gene="TTK"
/note="synonyms: MPS1L1, PYT, ESK"
/db_xref="LocusID:7272"
/db_xref="MIM:604092"
77. .2650
/codon_start=1
/product="TTK protein kinase"
/protein_id="AAH32858.1"
/db_xref="GI:23271249"
/db_xref="LocusID:7272"
/translation="MESEDLGRELITDSIMNKVRDIKNKPKNEDLTDELSLNKISAD
TTDNGTGNQIMMANNPEDWLSLLKLEKNSVPLSDALLNKLIGRYSQAIEALPPDK
YQNESFARIQVFAELKAIQEPDDARDYFQMARANCKKFAFVHISFAQFELSQCNVK
KSKQLLQKAVERGAVPELEMLEIALNLQKQLLSEEEKNLASTVLTAEQSFSGS
LGHQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYRNSLRQTNKTQSCPFGRVPEV
NLLNSPCDVKTDDSVVPCFMKQTSRSECRDLVVPKSPGSGNDSCELRNLKSVQNSH
FKEPLVSDKESSELIITDSITLKNKTESLLAKLEETKEYQEPVEPESNQKQWSKRRK
SECINQNPAASSNHQIPELARKVNTQKHTTTEQPVFVSQSPPISTSKWFDPKSI
CKTPSSNTLDDYMSCFRTPVKNDFPACQLSTPYQPCFQQQHQLATPLQNLQV
LASSANECISVKGRISYILKQIGSGSSKVFQVLEKQIYAIKYVNLEEDNQTLN
SYNEIAYLNKLQHSDKIIRLYDEITDQIYVMMECIGNIDLSWLKRSIDPWER
KSYWKMLEAVHTIHQHIVHSDLEKPFANFLIVDMLKIDFGIANQMOPDTTSVVKDS
QVGTVMPPPEAKDMSSRENGSKSKISPKSDVWSLGLIYMTYGTPTFOQIINQ
ISKLHAIIDPNHEIEFPDIPEKDLQVLKCLKRPKQKRIPELLAHPYVQIOTHPV
NQMAKGTTEEMKVLQVLGLNSPILKAATLYEHYSGESHSSSSSKTFEKKRGK
K"

misc_feature
1649. .2449
/note="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smart00220"

ORIGIN
Query Match 75.4%; Score 2914; DB 9; Length 2980;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2931; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
QY 925 GAAACAGTTTGGGTTCCATCTTTTTCATTTCCCGAGTCAGTTTCTGTAGAAATGGAAT 984
Db 24 GAAGGAGTTTGGGTTCCATCTTTTTCATTTCCCGAGTCAGTTTCTGTAGAAATGGAAT 83
QY 985 CCGAGGATTTAAGTGGCAGAGAATTGACAAATGATCCATAATGACAAAGTGAGAGACA 1044
Db 84 CCGAGGATTTAAGTGGCAGAGAATTGACAAATGATCCATAATGACAAAGTGAGAGACA 143
QY 1045 TTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTTGAATAAAATTTCTG 1104
Db 144 TTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTTGAATAAAATTTCTG 203

QY 1105 CTGATACTACAGATAAAGTCTCGGAACTGTTAAACCAAAATTTATGATGATGGCAAAACCAACCCAG 1164
Dd
204 CTGATACTACAGATAAAGTCTCGGAACTGTTAAACCAAAATTTATGATGATGGCAAAACCAACCCAG 263
QY 1165 AGGACTGGTTGAGTTGTTGCTCAAACTAGAGAAAAACAGTGTTCGGCTAAAGTATGCTC 1224
Dd 264 AGGACTGGTTGAGTTGTTGCTCAAACTAGAGAAAAACAGTGTTCGGCTAAAGTATGCTC 323
QY 1225 TTTTAAATAAATTGATTGGTGGTTACAGTCAAGCAATTGGAAGCGCTTCCCCCAAGATAAAT 1284
Dd 324 TTTTAAATAAATTGATTGGTGGTTACAGTCAAGCAATTGGAAGCGCTTCCCCCAAGATAAAT 383
QY 1285 ATGGCCAAAATGAGAGTTTGTGCTAGAAATCAAGTGAGATTTGCTGAATTTAAAGCTATTTC 1344
Dd 384 ATGGCCAAAATGAGAGTTTGTGCTAGAAATCAAGTGAGATTTGCTGAATTTAAAGCTATTTC 443
QY 1345 AAGAGCCAGATGATGCACGTGACTACTTTCAAAATGGCCAGAGCAAACTGCAAGAAATTTG 1404
Dd 444 AAGAGCCAGATGATGCACGTGACTACTTTCAAAATGGCCAGAGCAAACTGCAAGAAATTTG 503
QY 1405 CTTTGTGTTCAATATCTTTTGCACAAATTTGAACGTGTCAACAGTAAATGTCAAAAAAGTA 1464
Dd 504 CTTTGTGTTCAATATCTTTTGCACAAATTTGAACGTGTCAACAGTAAATGTCAAAAAAGTA 563
QY 1465 AACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGAAATGCTGGAATTTG 1524
Dd 564 AACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGAAATGCTGGAATTTG 623
QY 1525 CCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTCCAGAGGAGGAAAAAGAAAT 1584
Dd 624 CCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTCCAGAGGAGGAAAAAGAAAT 683
QY 1585 TATCAGCATCTACGGTATTAACTGCCCCAAGAAATCATTTTCCGGTTCACCTGGGCAATTTAC 1644
Dd 684 TATCAGCATCTACGGTATTAACTGCCCCAAGAAATCATTTTCCGGTTCACCTGGGCAATTTAC 743
QY 1645 AGAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTTTTATATG 1704
Dd 744 AGAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTTTTATATG 803
QY 1705 GAGAGAACATGCCACCAACAAGATGCAAGAAATAGGTTACCGGAATTCATTGAGACAAACTA 1764
Dd 804 GAGAGAACATGCCACCAACAAGATGCAAGAAATAGGTTACCGGAATTCATTGAGACAAACTA 863
QY 1765 ACAAACTAAACAGTCATGCCCATTTTGAAGAGTCCCAGTTAACTTCTTAAATAGCCAG 1824
Dd 864 ACAAACTAAACAGTCATGCCCATTTTGAAGAGTCCCAGTTAACTTCTTAAATAGCCAG 923
QY 1825 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTAAAGAAACAACTCTA 1884
Dd 924 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTAAAGAAACAACTCTA 983
QY 1885 GATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1944
Dd 984 GATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1043
QY 1945 AATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACCTCTGTTGTCAGATG 2004
Dd 1044 AATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACCTCTGTTGTCAGATG 1103
QY 2005 AAAAGAGTTCTGAACCTTATTATTACTGATTAACCTCAAGAAATGAAACGGAATCAA 2064
Dd 1104 AAAAGAGTTCTGAACCTTATTATTACTGATTAACCTCAAGAAATGAAACGGAATCAA 1163
QY 2065 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACCGAGGTTCCAGAGAGTA 2124
Dd 1164 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACCGAGGTTCCAGAGAGTA 1223
QY 2125 ACCAGAAACAGTGGCAACCTAAGAGAAAGTCAAGAGTGTATTAAACCAAGATCCTGCTGCAT 2184
Dd 1224 ACCAGAAACAGTGGCAACCTAAGAGAAAGTCAAGAGTGTATTAAACCAAGATCCTGCTGCAT 1283
QY 2185 CTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAAAAAGTTAATACAGAGCAGAAACATA 2244

Dd 1284 CTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAAAAAGTTAATACAGAGCAGAAACATA 1343
QY 2245 CCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAACAGTCAACCACCAATATCAACATCTA 2304
Dd 1344 CCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAACAGTCAACCACCAATATCAACATCTA 1403
QY 2305 AATGGTTTGACCCAAAATCTATTTGTAAAGACCAAGCAGCAATACCTTGGATGATTACA 2364
Dd 1404 AATGGTTTGACCCAAAATCTATTTGTAAAGACCAAGCAGCAATACCTTGGATGATTACA 1463
QY 2365 TGAGCTGTTTGTAGAACTCCAGTTGTAAAGAATGACTTTCACCTGCTTGTCAAGTTGTCAA 2424
Dd 1464 TGAGCTGTTTGTAGAACTCCAGTTGTAAAGAATGACTTTCACCTGCTTGTCAAGTTGTCAA 1523
QY 2425 CACCTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCATCAATACTTGCACCTCCAC 2484
Dd 1524 CACCTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCATCAATACTTGCACCTCCAC 1583
QY 2485 TTTCAAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAAAATGAATGCAATTTGGTTAAAGGAA 2544
Dd 1584 TTTCAAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAAAATGAATGCAATTTGGTTAAAGGAA 1643
QY 2545 GAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAGGTTTCAAGCAAGGTATTTTCAGGTTG 2604
Dd 1644 GAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAGGTTTCAAGCAAGGTATTTTCAGGTTG 1703
QY 2605 TAAATGAAAAGAAACAGATATATGCTATATAAATATGTGAACCTTAGAAGAACAGATAACC 2664
Dd 1704 TAAATGAAAAGAAACAGATATATGCTATATAAATATGTGAACCTTAGAAGAACAGATAACC 1763
QY 2665 AAACTCTTGATAGTTACCGGAACGAAATAGCTTATTTGAATATACTACAACACACAGTG 2724
Dd 1764 AAACTCTTGATAGTTACCGGAACGAAATAGCTTATTTGAATATACTACAACACACAGTG 1823
QY 2725 ATAAGATCATCCGACTTTATGATTATGAAATCAGGACAGTACATCTACATGTTAATGG 2784
Dd 1824 ATAAGATCATCCGACTTTATGATTATGAAATCAGGACAGTACATCTACATGTTAATGG 1883
QY 2785 AGTGTGAAATATTGATCTCTTAATAGTTGGCTTAAAAAGAAAAATCCATTCATCCATGGG 2844
Dd 1884 AGTGTGAAATATTGATCTCTTAATAGTTGGCTTAAAAAGAAAAATCCATTCATCCATGGG 1943
QY 2845 AACGAAAGAGTTACTGGAATAAATATGTTAGAGGACAGTTCAACAATCCATCAACATGGCA 2904
Dd 1944 AACGAAAGAGTTACTGGAATAAATATGTTAGAGGACAGTTCAACAATCCATCAACATGGCA 2003
QY 2905 TTGTTACAGTGATCTTAAACCCAGCTAACTTTCTGATAGTTGATGGAATGCTAAAGCTAA 2964
Dd 2004 TTGTTACAGTGATCTTAAACCCAGCTAACTTTCTGATAGTTGATGGAATGCTAAAGCTAA 2063
QY 2965 TTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGATACAACAAGTGTGTTAAAGATTCTC 3024
Dd 2064 TTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGATACAACAAGTGTGTTAAAGATTCTC 2123
QY 3025 AGGTTGGCACAGTTAATATATATGCCACCAGAGCAATCAAGATATGTTCTTCTCCAGAG 3084
Dd 2124 AGGTTGGCACAGTTAATATATATGCCACCAGAGCAATCAAGATATGTTCTTCTCCAGAG 2183
QY 3085 AGAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAAGTGTGTTTGGTCTTAGGATGTA 3144
Dd 2184 AGAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAAGTGTGTTTGGTCTTAGGATGTA 2243
QY 3145 TTTTGTACTATATGACTTACGGGAAAAACACCAATTTTCAGCAGATAATTAATCAGATTCTA 3204
Dd 2244 TTTTGTACTATATGACTTACGGGAAAAACACCAATTTTCAGCAGATAATTAATCAGATTCTA 2303
QY 3205 AATTACATGCCATAATTGATCCTTAATCATGAAATTTCCCGATATTTCCAGAGAAAG 3264
Dd 2304 AATTACATGCCATAATTGATCCTTAATCATGAAATTTCCCGATATTTCCAGAGAAAG 2363
QY 3265 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAACAGAGGATATCCATTTC 3324

Db 2364 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGGACCCAAACAGAGGATATCCATT 2423

QY 3325 CTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAACTCATCCAGTTAACCAATGGCCA 3384

Db 2424 CTGAGCTCCTGGCTCATCCATATGTTCAAAATTCAAACTCATCCAGTTAACCAATGGCCA 2483

QY 3385 AGGGAACCACTGAAGAAATGAATATGTTCTGGGCCAACTTGTGTTGCTGAATTCCTTA 3444

Db 2484 AGGGAACCACTGAAGAAATGAATATGTTCTGGGCCAACTTGTGTTGCTGAATTCCTTA 2543

QY 3445 ACTCCATTTTGAAGCTGCTAAACCTTTATATGAACACTATAGTGGTGGTGAAGTCATA 3504

Db 2544 ACTCCATTTTGAAGCTGCTAAACCTTTATATGAACACTATAGTGGTGGTGAAGTCATA 2603

QY 3505 ATTCTTCATCCTCCAAGACTTTTGAAGAAACCAAGGGGAAACCAATGATTGCAGTTATTC 3564

Db 2604 ATTCTTCATCCTCCAAGACTTTTGAAGAAACCAAGGGGAAACCAATGATTGCAGTTATTC 2663

QY 3565 GTAAATGTCAGATAGGAGGTATATAAATATATTGGACTGTTTATCTCTTGAATCCCTGTGGA 3624

Db 2664 GTAAATGTCAGATAGGAGGTATATAAATATATTGGACTGTTTATCTCTTGAATCCCTGTGGA 2723

QY 3625 AATCTACATTTGAAGACAACATCACTCTGAAGTGTATATCAGCAACCAAAATTCAGTGAGA 3684

Db 2724 AATCTACATTTGAAGACAACATCACTCTGAAGTGTATATCAGCAACCAAAATTCAGT-AGA 2782

QY 3685 TTATCTTTTAAAGAAACCTGTAAACCACTTATGCACTTATGCACTGTATATATTGTAGA 3744

Db 2783 TTATCTTTTAAAGAAACCTGTAAACCACTTATGCACTTATGCACTGTATATATTGTAGA 2842

QY 3745 CTTGTTTCTCTGTTTATGCTCTGTTGTAATCTACTTGACATCATTTTACTCTTGGAAAT 3804

Db 2843 CTTGTTTCTCTGTTTATGCTCTGTTGTAATCTACTTGACATCATTTTACTCTTGGAAAT 2902

QY 3805 AGTGGTGGATAGCAAGTATATCTTAAACCACTTGTAAATAAGTTTGTGGCTAAAT 3864

Db 2903 AGTGGTGGATAGCAAGTATATCTTAAACCACTTGTAAATAAGTTTGTGGCTAAAT 2962

QY 3865 GA 3866

Db 2963 GA 2964

RESULT 4

AX775733

LOCUS AX775733 2975 bp mRNA linear PAT 14-JUL-2003

DEFINITION Sequence 3 from Patent WO03048202.

ACCESSION AX775733

VERSION AX775733.1 GI:32693451

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Matsuda,A. and Muramatsu,S.

TITLE NF-kB activating Gene

JOURNAL Patent: WO 03048202-A 3 12-JUN-2003;

Asahi Kasei Kabushiki Kaisha (JP)

FEATURES

source

1. .2975

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

84. .2657

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE11566.1"

/db_xref="GI:32693452"

/translation="MESEDLSGRELIDSIMNKVRDINKFKNEDLTDELSLNKISAD

TTDSNGTVNQIMMANNPEDWLSLLKLEKNSVPLSDALNKLIGRYSQATEALPPDK

YQNESFARIQVRFAELKAIQEPDDARDYFQMARANCKFAFVHISFAQFELSQGNVK

KSKQLLQKAVERGAVPLEMLEIALRNLNQKQLLSEBEKNLSASTVLTAESFSGS

LGHLQNRNNSCDSRGOTTAKRFLYGENMPPQDAEIGYRNSLRQTNKTKQSCPFGRVPV

NLNSPDCDVKTDDSVVPCFMKQRTSECDLVVPGSPGNSDSCELNKLKSVQNSH

FKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQEPVESPKNQWQSKK

SECINQNPAASSNHQIPELARKVNTQKHITFEQPVFSKQSPPISTSKWFDPKSI

CKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQAPCFQOQHQHILATPLQNLQV

LASSANECISVKGRIYSILKQIGSGSSKVFQVNEKKQIYAIKYVNEEADNQTLD

SYRNEIAYLNKLOHSDKIIRLYDYEITDQIYIMVMECNIDLNSWLKXKSIDPWER

KSYWKNMLEAVHTTHQGHIVHSDLKPNFLIVDGMKLIDFGIANQMOPDTTSVVDKS

QVGTVNYMPPPEAIKDMSSSRENGKSKSPKSDVSWSLGCLILYMTYGTPTFQIINQ

ISKLHAIIDPNHEIFPDIPEKDLQVLKCLKRDPKRISIPPELLAHYPVQIOHPV

NQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLIEHYSGGSHNSSSKTTFEKRKG

K"

ORIGIN

Query Match 75.3%; Score 2909.2; DB 6; Length 2975;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 2928; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 925 GAAACACAGTTGGGTTCCATCTTTTCAATTTCCCAAGTGCAGTTTCTGTAGAAATGGAAT 984

Db 31 GAAAGGAGTTGGGTTCCATCTTTTCAATTTCCCAAGTGCAGTTTCTGTAGAAATGGAAT 90

QY 985 CCGAGGATTTAAGTGGCAGAGAATTGACAAATTGATTTCCATAATGAACAAAGTGAGAGACA 1044

Db 91 CCGAGGATTTAAGTGGCAGAGAATTGACAAATTGATTTCCATAATGAACAAAGTGAGAGACA 150

QY 1045 TTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAAATTTCTG 1104

Db 151 TTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAAATTTCTG 210

QY 1105 CTGATACTACAGATAAATCTCGGAACTGTAAACCAAAATATGATGATGCGCAACCAACCCAG 1164

Db 211 CTGATACTACAGATAAATCTCGGAACTGTAAACCAAAATATGATGATGCGCAACCAACCCAG 270

QY 1165 AGGACTGGTTGAGTTTGTGTTGCTCAAACTAGAGAAAAACAGTGTCCGCTAAGTGATGCTC 1224

Db 271 AGGACTGGTTGAGTTTGTGTTGCTCAAACTAGAGAAAAACAGTGTCCGCTAAGTGATGCTC 330

QY 1225 TTTTAAATAAATTTGATTGGTGGTTACAGTCAAGCAATTTGAAGCGCTTCCCCCAGATAAAT 1284

Db 331 TTTTAAATAAATTTGATTGGTGGTTACAGTCAAGCAATTTGAAGCGCTTCCCCCAGATAAAT 390

QY 1285 ATGGCCAAAATGAGAGTTTGTCTAGAAATTCAGTCAAGTGAATTTGCTGAATTTAAAGCTATT 1344

Db 391 ATGGCCAAAATGAGAGTTTGTCTAGAAATTCAGTCAAGTGAATTTGCTGAATTTAAAGCTATT 450

QY 1345 AAGAGCCAGATGATGACGCTGACTACTTTCAATGGCCAGAGCAAACTGCAAGAAATTTG 1404

Db 451 AAGAGCCAGATGATGACGCTGACTACTTTCAATGGCCAGAGCAAACTGCAAGAAATTTG 510

QY 1405 CTTTGTGTTTCATATATCTTTTGCACAAATTTGAAGTGTCAAGGTAATGTCAAAAAAGTA 1464

Db 511 CTTTGTGTTTCATATATCTTTTGCACAAATTTGAAGTGTCAAGGTAATGTCAAAAAAGTA 570

QY 1465 AACAACTTCTTCAAAAAGCTGTAGAACGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1524

Db 571 AACAACTTCTTCAAAAAGCTGTAGAACGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 630

QY 1525 CCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGCTTTTCAGAGGAGGAGAGAGAGAGAG 1584

Db 631 CCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGCTTTTCAGAGGAGGAGAGAGAGAGAG 690

QY 1585 TATCAGCATCTACGGTATTAACTGCCCAAGATCATTTTCCGGTTCACCTTGGGCATTTAC 1644

Db 691 TATCAGCATCTACGGTATTAACTGCCCAAGATCATTTTCCGGTTCACCTTGGGCATTTAC 750

QY 1645 AGAATAGGAACAACACAGTTGTGATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1704

Db 751 AGAATAGGAACAACACAGTTGTGATTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 810

QY 1705 GAGAGAACATGCCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGACAAACTA 1764

Db 811 GAGAGAACATGCCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGACAAACTA 870

QY 1765 ACAAACATAACAGTCAATGCCATTTGGAGAGTCCAGTTAAACCTTCTTAATAGCCAG 1824
Db |||||
QY 871 ACAAACATAACAGTCAATGCCATTTGGAGAGTCCAGTTAAACCTTCTTAATAGCCAG 930
Db |||||
QY 1825 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGTTTATGAAAAGACAAACCTCTA 1884
Db |||||
QY 931 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGTTTATGAAAAGACAAACCTCTA 990
Db |||||
QY 1885 GATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1944
Db |||||
QY 991 GATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1050
Db |||||
QY 1945 AATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCAATTTCAAGGAACCTCTGTGTGCAGATG 2004
Db |||||
QY 1051 AATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCAATTTCAAGGAACCTCTGTGTGCAGATG 1110
Db |||||
QY 2005 AAAAGAGTTCGAACCTTATTATTACTGATTTCAATAACCTTGAAGAAATAAAACGGAATCAA 2064
Db |||||
QY 1111 AAAAGAGTTCGAACCTTATTATTACTGATTTCAATAACCTTGAAGAAATAAAACGGAATCAA 1170
Db |||||
QY 2065 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACAGAGGTTCCAGAGAGTA 2124
Db |||||
QY 1171 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACAGAGGTTCCAGAGAGTA 1230
Db |||||
QY 2125 ACCAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGTGATTAACAGAAATCCTGTCTGCAT 2184
Db |||||
QY 1231 ACCAGAAACAGTGGCAATCTAAGAGAAAGTCAGAGTGATTAACAGAAATCCTGTCTGCAT 1290
Db |||||
QY 2185 CTTCAAAATCACTGGCAGATTCGGGAGTTAGCCGAAAGTTAAATACAGAGCAAGAAACATA 2244
Db |||||
QY 1291 CTTCAAAATCACTGGCAGATTCGGGAGTTAGCCGAAAGTTAAATACAGAGCAAGAAACATA 1350
Db |||||
QY 2245 CCACCTTTTGAGCAACCTGTCTTTTCAAGTTTCAAAACAGTCACCAACCAATATCAACATCTA 2304
Db |||||
QY 1351 CCACCTTTTGAGCAACCTGTCTTTTCAAGTTTCAAAACAGTCACCAACCAATATCAACATCTA 1410
Db |||||
QY 2305 AATGGTTTGACCCAAATCTATTGTGAAGACACCAAGCAGCAATACCTTGGATGATTACA 2364
Db |||||
QY 1411 AATGGTTTGACCCAAATCTATTGTGAAGACACCAAGCAGCAATACCTTGGATGATTACA 1470
Db |||||
QY 2365 TGAGCTGTTTGAACCTCCAGTTGTAAAGAATGACTTTTCCACCTGCTTGTCTAGTTGTCAA 2424
Db |||||
QY 1471 TGAGCTGTTTGAACCTCCAGTTGTAAAGAATGACTTTTCCACCTGCTTGTCTAGTTGTCAA 1530
Db |||||
QY 2425 CACCTTATGGCCAACTGCCTGTTTCCAGCAGCAACAGCATCAAAATACTTGGCACTCCAC 2484
Db |||||
QY 1531 CACCTTATGGCCAACTGCCTGTTTCCAGCAGCAACAGCATCAAAATACTTGGCACTCCAC 1590
Db |||||
QY 2485 TTCAAAATTTACAGGTTTTCAGCATCTTCTTCCAGCAAAATGAATGCATTTCCGTTAAAGGAA 2544
Db |||||
QY 1591 TTCAAAATTTACAGGTTTTCAGCATCTTCTTCCAGCAAAATGAATGCATTTCCGTTAAAGGAA 1650
Db |||||
QY 2545 GAATTTATTTCCATATTAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTCAGGTGT 2604
Db |||||
QY 1651 GAATTTATTTCCATTTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTCAGGTGT 1710
Db |||||
QY 2605 TAAATGAAAAGAAACAGATATATGCTATAAAATATGTGAACTTAGAAGAGCAGATAACC 2664
Db |||||
QY 1711 TAAATGAAAAGAAACAGATATATGCTATAAAATATGTGAACTTAGAAGAGCAGATAACC 1770
Db |||||
QY 2665 AAACCTCTTGATAGTTACCGGAACGAAATAGCTTATTGAAATAAACTACAACAAACAGATG 2724
Db |||||
QY 1771 AAACCTCTTGATAGTTACCGGAACGAAATAGCTTATTGAAATAAACTACAACAAACAGATG 1830
Db |||||
QY 2725 ATAAGATCATCCGACTTTATGATTAATGAATACCGGACCAAGTACATCTACATGGTAATGG 2784
Db |||||
QY 1831 ATAAGATCATCCGACTTTATGATTAATGAATACCGGACCAAGTACATCTACATGGTAATGG 1890
Db |||||
QY 2785 AGTGTGGAATATTGATCTTAATAGTTGGCTTAAAAAGAAAAAATCCATTTGATCCATGGG 2844
Db |||||
QY 1891 AGTGTGGAATATTGATCTTAATAGTTGGCTTAAAAAGAAAAAATCCATTTGATCCATGGG 1950
Db |||||

QY 2845 AACGCAAGAGTTACTGGAAAAATATGTTAGAGGCGAGTTACACAATCCATCAACATGGCA 2904
Db |||||
QY 1951 AACGCAAGAGTTACTGGAAAAATATGTTAGAGGCGAGTTACACAATCCATCAACATGGCA 2010
Db |||||
QY 2905 TTGTTTACAGTGATCTTAAACCAGCTAACTTTCTGATAGTTGGAATGCTAAAGCTAA 2964
Db |||||
QY 2011 TTGTTTACAGTGATCTTAAACCAGCTAACTTTCTGATAGTTGGAATGCTAAAGCTAA 2070
Db |||||
QY 2965 TTGATTTTGGGATTGCAAAACCAAAATGCAACCAGATACAAAGTGTGTTAAAGATTCTC 3024
Db |||||
QY 2071 TTGATTTTGGGATTGCAAAACCAAAATGCAACCAGATACAAAGTGTGTTAAAGATTCTC 2130
Db |||||
QY 3025 AGGTTGGCACAGTTAAATATATGCCACCAAGCAATCAAGATATGTCTTCTCCAGAG 3084
Db |||||
QY 2131 AGGTTGGCACAGTTAAATATATGCCACCAAGCAATCAAGATATGTCTTCTCCAGAG 2190
Db |||||
QY 3085 AGAATGGGAAATCTAAAGTCAAGATAAAGCCCCAAAAAGTGTGTTTGGTCCCTTAGGATGTA 3144
Db |||||
QY 2191 AGAATGGGAAATCTAAAGTCAAGATAAAGCCCCAAAAAGTGTGTTTGGTCCCTTAGGATGTA 2250
Db |||||
QY 3145 TTTTGTACTATATGACTTACGGGAAAAACACCATTTTCCAGAGATAAATAATCAGATTCTA 3204
Db |||||
QY 2251 TTTTGTACTATATGACTTACGGGAAAAACACCATTTTCCAGAGATAAATAATCAGATTCTA 2310
Db |||||
QY 3205 AATTACATGCCATAATTGATCTTAATCATGAAATTTGAAATTTCCCGATATTTCCAGAGAAAG 3264
Db |||||
QY 2311 AATTACATGCCATAATTGATCTTAATCATGAAATTTGAAATTTCCCGATATTTCCAGAGAAAG 2370
Db |||||
QY 3265 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGGACCCAAAAACAGAGGATATCCATTTC 3324
Db |||||
QY 2371 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGGACCCAAAAACAGAGGATATCCATTTC 2430
Db |||||
QY 3325 CTGAGCTCCTGGCTCATCCATATGTTCAAATTTCAAACCTCATCCAGTTAAACCAATGGCCA 3384
Db |||||
QY 2431 CTGAGCTCCTGGCTCATCCCTATGTTCAAATTTCAAACCTCATCCAGTTAAACCAATGGCCA 2490
Db |||||
QY 3385 AGGGAACCACTGAAGAAATGAAATATGTTCTGGGCCAACTTGTGGTCTGGAATTTCTCCTA 3444
Db |||||
QY 2491 AGGGAACCACTGAAGAAATGAAATATGTTCTGGGCCAACTTGTGGTCTGGAATTTCTCCTA 2550
Db |||||
QY 3445 ACTCCATTTTGAAGCTGCTAAACCTTTATATGAACACTATAGTGGTGGTGAAGTCTATA 3504
Db |||||
QY 2551 ACTCCATTTTGAAGCTGCTAAACCTTTATATGAACACTATAGTGGTGGTGAAGTCTATA 2610
Db |||||
QY 3505 ATTCTTCATCCTCCAAGACTTTTGAAGAAAGGGGAAAAAATGATTTGCAGTTATTTC 3564
Db |||||
QY 2611 ATTCTTCATCCTCCAAGACTTTTGAAGAAAGGGGAAAAAATGATTTGCAGTTATTTC 2670
Db |||||
QY 3565 GTAATGTGAGATAGGAGGTATAAAATATATTTGGACTGTTTACTCTTGAATCCCTGTGGA 3624
Db |||||
QY 2671 GTAATGTCAAAATACCACCTATAAAATATATTTGGACTGTTTACTCTTGAATCCCTGTGGA 2730
Db |||||
QY 3625 AATCTACATTTGAAGACAAACATCACTCTGAAAGTGTATATCAGCAAAAAAATTCAGTGAGA 3684
Db |||||
QY 2731 AATCTACATTTGAAGACAAACATCACTCTGAAAGTGTATATCAGCAAAAAAATTCAGT-AGA 2789
Db |||||
QY 3685 TTATCTTTAAAGAAAACTGTAAAAATAGCAACCACTTATGGCACTGTATATATTGTAGA 3744
Db |||||
QY 2790 TTATCTTTAAAGAAAACTGTAAAAATAGCAACCACTTATGGTACTGTATATATTGTAGA 2849
Db |||||
QY 3745 CTTGTTTCTCTGTTTATGCTCTTGTGTAATCTACTTGACATCATTTTACTCTTGGAAAT 3804
Db |||||
QY 2850 CTTGTTTCTCTGTTTATGCTCTTGTGTAATCTACTTGACATCATTTTACTCTTGGAAAT 2909
Db |||||
QY 3805 AGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTGTAAAAAAGTTTGTGGCTAAAAAT 3864
Db |||||
QY 2910 AGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTGTAAAAAAGTTTGTGGCTAAAAAT 2969
Db |||||
QY 3865 GA 3866
Db |||||
QY 2970 GA 2971
Db |||||

RESULT 5
BC000633
LOCUS
DEFINITION
Homo sapiens TTK protein kinase, mRNA (cdna clone MGC:865
IMAGE:3343925), complete cds.
ACCESSION
BC000633
VERSION
BC000633.2 GI:33990673
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2984)
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2984)
Strausberg,R.
Direct Submission
Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 20, 2003 this sequence version replaced gi:12653696.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 5 Row: d Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 23308721.
Location/Qualifiers
1. .2984
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="MGC:865 IMAGE:3343925"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. .2984
/gene="TTK"
/note="synonyms: MPS1L1, PYT, ESK"
/db_xref="LocusID:7272"
/db_xref="MIM:604092"
75. .2648
/codon_start=1
/product="TTK protein kinase"
/protein_id="AAH00633.1"
/db_xref="GI:12653697"
/db_xref="LocusID:7272"
/translation="MESEDLSGRELTIDSIMNKVRDIKNFKVEDLTDELSLNKISAD
TTDNGTVNQIMMANNPEDWLSLLKLEKNSVPLSDALLNKLIGRYSQAIEALPPDK
YQNESFARIOVRFAELKAIOEPDDARDYFQMARANCKKFAFVHISFAQFELSQGNVK
XSKQLQKAVERGAAPLEMLEIALNLNQLKQLLSEEEKNLASASTVLTAESEFSGS
LGHQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYRNSLRQTNKTKQSCPFGRVPV
NLLNSPCDVKTDDSDVPCFMKQTSRSECDLVVPGKSPGSDNSCELNLKSVQNSH
FKEPLVSDKESSELIITDSITLKNKTESLLAKLEETKEYQEPEVPSNQKQWQSKRK
SECINQNPAASSNHQIPELARKVNTTEKHHTTFFQPVFSVSKSPPISTSKWDFPKSI
CKTPSSNTLDDYMSCFRTPVVKNDFFPACQLSTPYGQACFQQQHQILATPLQLIQV
LASSANECISVKGRISILKQIGSGSSKVFQVNLKXQIYAIKYVNEEADNQTLD
SYRNEIAYLNKLQOHSDKIIRLYDEITDQIYVMECIGNIDLSWLKXKSIDPWER
KSYWKNMLEAVHTIHOHGIVHSDLPANFLIVDMLKLIIDFGIANQMOPDTTSSVVDK
QVGTVNYMPPBEAKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYKTPFQOILNQ
ISKLHAIIDPNHEIEFPDIPEKLDQVLKCLKRPKQISIPPELLAHYPYVQIQTHPV
NQMAKGTTEEMKVVLGQLVGLNSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK
K"
1647. .2447
/note="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smart00220"
ORIGIN
Query Match 75.3%; Score 2909.2; DB 9; Length 2984;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2928; Conservative 0; Mismatches 13; Indels 1; Gaps 1;
QY 925 GAAACACAGTTTGGGTTCCATCTTTTTCATTTCCCGCAGTGCAGTTTCTGTAGAAATGGAAT 984
DB 22 GAAAGGAGTTTGGGTTCCATCTTTTTCATTTCCCGCAGCGCAGCTTTCTGTAGAAATGGAAT 81
QY 985 CCGAGGATTTAAGTGGCAGAGAATTGACAAATGATTCCATAATGAACAAAGTGAGAGACA 1044
DB 82 CCGAGGATTTAAGTGGCAGAGAATTGACAAATGATTCCATAATGAACAAAGTGAGAGACA 141
QY 1045 TTTAAAAATAAGTTTAAAAATAAGAACCTTACTGATGAACCTTAAAGCTTGAATAAAATTTCTG 1104
DB 142 TTTAAAAATAAGTTTAAAAATAAGAACCTTACTGATGAACCTTAAAGCTTGAATAAAATTTCTG 201
QY 1105 CTGATACTACAGATAACTCGGGAACTGTGTAACCAAAATTTATGATGATGGCAACCAACCCAG 1164
DB 202 CTGATACTACAGATAACTCGGGAACTGTGTAACCAAAATTTATGATGATGGCAACCAACCCAG 261
QY 1165 AGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCGGCTAAGTGATGCTC 1224
DB 262 AGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCGGCTAAGTGATGCTC 321
QY 1225 TTTTAAATAAATTTGATTTGCTGTACAGTCAAGCAATTTGAAGCGCTTCCCCCAGATAAAT 1284
DB 322 TTTTAAATAAATTTGATTTGCTGTACAGTCAAGCAATTTGAAGCGCTTCCCCCAGATAAAT 381
QY 1285 ATGCCAAAATGAGAGTTTGTGCTAGAAATTCAGAGTGAATTTTGTGAATTTAAAGCTATTTC 1344
DB 382 ATGCCAAAATGAGAGTTTGTGCTAGAAATTCAGAGTGAATTTTGTGAATTTAAAGCTATTTC 441
QY 1345 AAGAGCCAGATGATGCAGTGAAGTACTTTTCAAAATGGCCAGAGCAAACTGCAAGAAATTTG 1404

Db 442 AAGAGCCAGATGATGCACGTGACTACTTTCAAATGGCCAGAGCAAACTGCAAGAAATTG 501

Qy 1405 CTTTGTTCATATATCTTTTGCACAATTTGAACCTGTCAACAAGTAAATGTCAAAAAAGTA 1464

Db 502 CTTTGTTCATATATCTTTTGCACAATTTGAACCTGTCAACAAGTAAATGTCAAAAAAGTA 561

Qy 1465 AACAACTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCACATAGAAAATGCTGGAATG 1524

Db 562 AACAACTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCACATAGAAAATGCTGGAATG 621

Qy 1525 CCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTCAGAGGAGGAAAAAGACAAAT 1584

Db 622 CCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTCAGAGGAGGAAAAAGACAAAT 681

Qy 1585 TATCAGCATCTACGGTATTAACCTGCCCCAAGAAATCATTTCCGGTTCACCTGGGCAATTAC 1644

Db 682 TATCAGCATCTACGGTATTAACCTGCCCCAAGAAATCATTTCCGGTTCACCTGGGCAATTAC 741

Qy 1645 AGAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTTTATATG 1704

Db 742 AGAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTTTATATG 801

Qy 1705 GAGAGAACATGCCACACACAAGATGCAGAAATAGTTTACCGGAATTCATTGAGACAACTA 1764

Db 802 GAGAGAACATGCCACACACAAGATGCAGAAATAGTTTACCGGAATTCATTGAGACAACTA 861

Qy 1765 AAAAACTAAACAGTCATGCCCATTTTGGAGAGTCCCAGTTAACCTTCTAAATAGCCAG 1824

Db 862 AAAAACTAAACAGTCATGCCCATTTTGGAGAGTCCCAGTTAACCTTCTAAATAGCCAG 921

Qy 1825 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGAAAAGACAAACCTCTA 1884

Db 922 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGAAAAGACAAACCTCTA 981

Qy 1885 GATCAGAAATGCCGAGATTTGGTTGTGCTTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1944

Db 982 GATCAGAAATGCCGAGATTTGGTTGTGCTTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1041

Qy 1945 AATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATTTCAAGGAACCTCTGTTGTGATG 2004

Db 1042 AATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATTTCAAGGAACCTCTGTTGTGATG 1101

Qy 2005 AAAAGAGTTCTGAACCTATTATTACTGATTCATAACCTGAGAAATAAAAACGGAATCAA 2064

Db 1102 AAAAGAGTTCTGAACCTATTATTACTGATTCATAACCTGAGAAATAAAAACGGAATCAA 1161

Qy 2065 GTCTTCTAGCTAAATTAGAAGAAACTAAAGAGTATCAAGAACAGAGGTTCCAGAGAGTA 2124

Db 1162 GTCTTCTAGCTAAATTAGAAGAAACTAAAGAGTATCAAGAACAGAGGTTCCAGAGAGTA 1221

Qy 2125 ACCAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAGTGTATTAAACCAGAAATCCTGCTGCAT 2184

Db 1222 ACCAGAAAACAGTGGCAATCTAAGAGAAAGTCAGAGTGTATTAAACCAGAAATCCTGCTGCAT 1281

Qy 2185 CTTCAAAATCACTGGCAGATTCGGGAGTTAGCCCGAAAAGTTAATACAGAGCAGAAAACATA 2244

Db 1282 CTTCAAAATCACTGGCAGATTCGGGAGTTAGCCCGAAAAGTTAATACAGAGCAGAAAACATA 1341

Qy 2245 CCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAACAGTCACCAACCAATATCAACATCTA 2304

Db 1342 CCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAACAGTCACCAACCAATATCAACATCTA 1401

Qy 2305 AATGGTTTGACCCAAAATCTATTGTAAGACACCAAGCAGCAATACCTTGGATGATTACA 2364

Db 1402 AATGGTTTGACCCAAAATCTATTGTAAGACACCAAGCAGCAATACCTTGGATGATTACA 1461

Qy 2365 TGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTGATGTTCAA 2424

Db 1462 TGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTGATGTTCAA 1521

Qy 2425 CACCTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCATCAAAATACTTGGCCTCCAC 2484

Db 1522 CACCTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCATCAAAATACTTGGCCTCCAC 1581

Qy 2485 TTCAAAAATTTACAGGTTTTCAGCATCTTCTTCAGCAAAATGAATGCATTTCCGTTAAAGGAA 2544

Db 1582 TTCAAAAATTTACAGGTTTTCAGCATCTTCTTCAGCAAAATGAATGCATTTCCGTTAAAGGAA 1641

Qy 2545 GAATTTATTCATATTAAAGCAGATAGGAAGTGGAGGTTTCAAGCAAGGTATTTTCAGGTGT 2604

Db 1642 GAATTTATTCATTTTAAAGCAGATAGGAAGTGGAGGTTTCAAGCAAGGTATTTTCAGGTGT 1701

Qy 2605 TAAATGAAAAAGAAACAGATATATGCTATATAAATATGTGAACCTTAGAAGAAAGCAGATAACC 2664

Db 1702 TAAATGAAAAAGAAACAGATATATGCTATATAAATATGTGAACCTTAGAAGAAAGCAGATAACC 1761

Qy 2665 AAACCTCTTGATAGTTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACACACAGTG 2724

Db 1762 AAACCTCTTGATAGTTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACACACAGTG 1821

Qy 2725 ATAAGATCATCCGACTTATGATTATGAATCAGGACCCAGTACATCTACATGGTAATGG 2784

Db 1822 ATAAGATCATCCGACTTATGATTATGAATCAGGACCCAGTACATCTACATGGTAATGG 1881

Qy 2785 AGTGTGGAATATTGATCTTAAATAGTTGGCTTAAAAAGAAAAATCCATTCATCCATGGG 2844

Db 1882 AGTGTGGAATATTGATCTTAAATAGTTGGCTTAAAAAGAAAAATCCATTCATCCATGGG 1941

Qy 2845 AACGCAAGAGTTACTGGAAAAATATGTTAGAGCAGTTTCAACATCCATCAACATGGCA 2904

Db 1942 AACGCAAGAGTTACTGGAAAAATATGTTAGAGCAGTTTCAACATCCATCAACATGGCA 2001

Qy 2905 TTGTTCAAGTGTATCTTAAACCAGCTAACTTTCTGTATGATGTAAGTAAAGCTAA 2964

Db 2002 TTGTTCAAGTGTATCTTAAACCAGCTAACTTTCTGTATGATGTAAGTAAAGCTAA 2061

Qy 2965 TTGATTTTGGGATTGCAAAACCAATGCAACCCAGATACAACAAGTGTGTTTAAAGATTCTC 3024

Db 2062 TTGATTTTGGGATTGCAAAACCAATGCAACCCAGATACAACAAGTGTGTTTAAAGATTCTC 2121

Qy 3025 AGTTTGGCACAGTTAATATATATGCCACCAGAAAGCAATCAAAAGATATGTTTCTCCAGAG 3084

Db 2122 AGTTTGGCACAGTTAATATATATGCCACCAGAAAGCAATCAAAAGATATGTTTCTCCAGAG 2181

Qy 3085 AGAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAAGTGTGTTTGGTCTTAGGATGTA 3144

Db 2182 AGAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAAGTGTGTTTGGTCTTAGGATGTA 2241

Qy 3145 TTTTGTACTATATGACTTACGGGAAAAACACCATTTTCAGCAGATAAATAATCAGATTCTA 3204

Db 2242 TTTTGTACTATATGACTTACGGGAAAAACACCATTTTCAGCAGATAAATAATCAGATTCTA 2301

Qy 3205 AATTACATGCCATAATTGATCCTTAATCATGAATTTGAATTTCCCGATATTCAGAGAAAG 3264

Db 2302 AATTACATGCCATAATTGATCCTTAATCATGAATTTGAATTTCCCGATATTCAGAGAAAG 2361

Qy 3265 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGGACCCAAAACAGAGGATATCCATTTC 3324

Db 2362 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGGACCCAAAACAGAGGATATCCATTTC 2421

Qy 3325 CTGAGCTCCTGGCTCATCCATATGTTTCAAAATCAAACTCATCCAGTTAAACCAATGGCCA 3384

Db 2422 CTGAGCTCCTGGCTCATCCCTATGTTTCAAAATCAAACTCATCCAGTTAAACCAATGGCCA 2481

Qy 3385 AGGAAACCACTGAAGAAATGAATAATGTTCTGGGCCCAA CTGTTGGTCTGAAATTCCTCTA 3444

Db 2482 AGGAAACCACTGAAGAAATGAATAATGTTCTGGGCCCAA CTGTTGGTCTGAAATTCCTCTA 2541

Qy 3445 ACTCCATTTTGAAGCTGTAAAACTTTTATATGAACACTATAGTGGTGGTGAAGTCTATA 3504

Db 2542 ACTCCATTTTGAAGCTGTAAAACTTTTATATGAACACTATAGTGGTGGTGAAGTCTATA 2601

Qy 3505 ATTCTTCATCCTCCAAAGACTTTTGAAAAAAAGGGGAAAAAATGATTTGCAGTTATTTC 3564

Db 2602 ATTCTTCATCCTCCAAAGACTTTTGAAAAAAAGGGGAAAAAATGATTTGCAGTTATTTC 2661

Db 1321 |||||CGAAAAGTTAATACAGAGCAGAAACATACCACTTTTGAGCAACGCTGCTTTTCAGTTTCA 1380
QY 2277 AAACAGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTGAAGACA 2336
Db 1381 |||||AAACAGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTGAAGACA 1440
QY 2337 CCAAGCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAT 2396
Db 1441 CCAAGCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAT 1500
QY 2397 GACTTTCACCTGCTGCTGTCAGTTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAG 2456
Db 1501 GACTTTCACCTGCTGCTGTCAGTTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAG 1560
QY 2457 CAACAGCATCAAATACTTGCACCTCCACTTCAAAATTTTACAGGTTTGTAGCATCTTCTTCA 2516
Db 1561 CAACAGCATCAAATACTTGCACCTCCACTTCAAAATTTTACAGGTTTGTAGCATCTTCTTCA 1620
QY 2517 GCAAAATGAATGCATTCGGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGT 2576
Db 1621 GCAAAATGAATGCATTCGGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGT 1680
QY 2577 GGAGGTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAGAAACAGATATATGCTATAAAA 2636
Db 1681 GGAGGTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAGAAACAGATATATGCTATAAAA 1740
QY 2637 TATGTGAACCTTAGAAGAGCAGATAACCAACTCTTGTATAGTTACCGGAACGAAATAGCT 2696
Db 1741 TATGTGAACCTTAGAAGAGCAGATAACCAACTCTTGTATAGTTACCGGAACGAAATAGCT 1800
QY 2697 TATTTGAATAAACTACAACAAACACAGTGATAAGATCATCCGACTTTATGATGAATGAAATC 2756
Db 1801 TATTTGAATAAACTACAACAAACACAGTGATAAGATCATCCGACTTTATGATGAATGAAATC 1860
QY 2757 ACGGACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTT 2816
Db 1861 ACGGACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTT 1920
QY 2817 AAAAAAGAAAAATCCATTCATGCGGAACGCAAGAGTTACTGGAAAAATATGTTAGAG 2876
Db 1921 AAAAAAGAAAAATCCATTCATGCGGAACGCAAGAGTTACTGGAAAAATATGTTAGAG 1980
QY 2877 GCAGTTCACACAATCCATCAACATGGCATTGTTCCAGTGATCTTAAACCAGCTAATCTT 2936
Db 1981 GCAGTTCACACAATCCATCAACATGGCATTGTTCCAGTGATCTTAAACCAGCTAATCTT 2040
QY 2937 CTGATAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTGCAAAACCAATGCAACCA 2996
Db 2041 CTGATAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTGCAAAACCAATGCAACCA 2100
QY 2997 GATACAACAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATTATATGCCACCAGAA 3056
Db 2101 GATACAACAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATTATATGCCACCAGAA 2160
QY 3057 GCAATCAAAGATATGTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCC 3116
Db 2161 GCAATCAAAGATATGTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCC 2220
QY 3117 AAAAGTGATGTTTGGTCTTAGGATGTTATTTTGTACTATATGACTTACGGGAAAAACCA 3176
Db 2221 AAAAGTGATGTTTGGTCTTAGGATGTTATTTTGTACTATATGACTTACGGGAAAAACCA 2280
QY 3177 TTTTCAGCAGATAATTAATCAGATTTCTAAATTAATGATGCCATAATGATCCTAATCATGAA 3236
Db 2281 TTTTCAGCAGATAATTAATCAGATTTCTAAATTAATGATGCCATAATGATCCTAATCATGAA 2340
QY 3237 ATTGAATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAA 3296
Db 2341 ATTGAATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAA 2400
QY 3297 AGGACCCCAAAACAGAGGATATCCATTCTCGAGCTCTCGCTCATCCATATGTTTCAAAAT 3356

Db 2401 AGGACCCCAAAACAGAGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAAT 2460
QY 3357 CAAACTCATCCAGTTAACCAAAATGGCCAAAGGAACCACTGAAGAAATGAAATATGTTCTG 3416
Db 2461 CAAACTCATCCAGTTAACCAAAATGGCCAAAGGAACCACTGAAGAAATGAAATATGTTCTG 2520
QY 3417 GGCCAACTTGTGTTGGTCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTAAAACTTTATAT 3476
Db 2521 GGCCAACTTGTGTTGGTCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTAAAACTTTATAT 2580
QY 3477 GAACACTATAGTGGTGGTGAAGTGCATAATTTTTCATCCTCCAAGACTTTTGAAAAAAAA 3536
Db 2581 GAACACTATAGTGGTGGTGAAGTGCATAATTTTTCATCCTCCAAGACTTTTGAAAAAAAA 2640
QY 3537 AGGGAAAAAAATGATTTCAGTTATTTCGTAATGTCAGATAGGAGGTATATAAATATATTG 3596
Db 2641 AGGGAAAAAAATGATTTCAGTTATTTCGTAATGTCAGATAGGAGGTATATAAATATATTG 2700
QY 3597 GACTGTTTACTCTTGAATCCCTGTGGAAATCTGATTAATCTTAAAAAGAAAACTGAAAAATAGCAA 3716
Db 2701 GACTGTTTACTCTTGAATCCCTGTGGAAATCTGATTAATCTTAAAAAGAAAACTGAAAAATAGCAA 2819
QY 3717 CCACCTATGGCAGCTGTATATATTTGTAGACTGTTGTTCTCTGTTTATGCTCTTGTGTAAT 3776
Db 2820 CCACCTATGGCAGCTGTATATATTTGTAGACTGTTGTTCTCTGTTTATGCTCTTGTGTAAT 2879
QY 3777 CTACTTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTCTAAAAAAC 3836
Db 2880 CTACTTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTCTAAAAAAC 2939
QY 3837 TTTGTAATAAAGTTTGTGGCTAAAATGA 3866
Db 2940 TTTGTAATAAAGTTTGTGGCTAAAATGA 2969

RESULT 7
AK095599
LOCUS AK095599 2419 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ38280 fis, clone FCBBF3005497, highly similar to DUAL SPECIFICITY PROTEIN KINASE TTK (EC 2.7.1.-).
ACCESSION AK095599
VERSION AK095599.1 GI:21754889
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T., Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiya, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Terashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuo, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2419)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of

QY 2735 CCGACTTTATGATTATGAATACCGGACCAGTACATCTACATGTTAATGGAGTGTGAAA 2794
|||||
Db 2338 CCGACTTTATGATTATGAATACCGGACCAGTACATCTACATGTTAATGGAGTGTGAAA 2397
|||||
QY 2795 TATTGATCTTAATAGTTGGCTT 2816
|||||
Db 2398 TATTGATCTTAATAGTTGGCTT 2419
|||||

RESULT 8
BC058851 2906 bp mRNA linear ROD 21-OCT-2003
LOCUS Mus musculus Ttk protein kinase, mRNA (cDNA clone MGC:60685
DEFINITION IMAGE:30023533), complete cds.

ACCESSION BC058851
VERSION BC058851.1 GI:37589052
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2906)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettaman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
PUBMED 22388257
2 (bases 1 to 2906)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Janet Rossant and Tilo Kunath (Samuel
Lunenfeld Research Institute, Canada)
cDNA Library Preparation: Yulan Piao and Minoru Ko (National
Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 111 Row: m Column: 18

This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6678462.
Location/Qualifiers

FEATURES
source 1..2906
/organism="Mus musculus"
/mol_type="mRNA"
/strain="B5/EGFP transgenic ICR mice"
/db_xref="taxon:10090"
/clone="MGC:60685 IMAGE:30023533"
/tissue_type="Trophoblast Stem Cell, mouse, 3.5 dpc"
/clone_lib="NIA Mouse Trophoblast Stem Cell cDNA Library
(Long)"
/lab_host="DH10B"
/note="Vector: pSPORT1"
gene 1..2906
/gene="Ttk"
/note="synonyms: PYT, esk, Esk1"
/db_xref="LocusID:22137"
/db_xref="MGI:1194921"

CDS 76..2568
/codon_start=1
/product="Ttk protein"
/protein_id="AAH58851.1"
/db_xref="GI:37589053"
/db_xref="LocusID:22137"
/translation="MEAEELIGSSVTIDSIMSKMRDIKNKINEDCTDELSLSKICADH
TETVNIQIMRVGNTPENWLNFLLLKLEKNSSPLNDDLLNKLIGRYSAIEVLPDPKYGN
ESFARIQVRLAELKAIQEPDDARDYFQMARENCKKFAFVHVSFAQFELSQGNLKKSEQ
LLHKAVETGAVPLQMLETAMRNHLQKQLPEEDKSVASATVLSAQEPFSSSLGNV
QNRISCESRGQAARVLYGENLPQDAEVRHQNPFQKTHAAKRSCPFGRVPVNLN
SPDFVVKTDSSAVTQLTTRKSGSPDRDAIILPGSRPRGSDSYELRGLKPIQTIYLKDSL
VSEKSSSELSDLIALKSKTDSLTLEETKPEIAERRPMQWQSTRKPECFVQNPAAF
APLRHVPDVTLKADKESPPISVPKWLDPKSACETPSSSSLDYMKCFKTPVKNDFPP
ACPSSTPYSQLARLQQQQQGLSTPLQSLQISGSSSINECISVNGRIYSILKQIDSGG
SSKVFQVLNEKKQINAIKYVNLEDADSTIESYRNEIAFLNKLQHQHSDKIIRLYDYEI
TEQIYIMVMCEGNI DLNWLKKSINPWERKSYWKNMVEAVHIIHQGIVHSDLKPA
NFVIVDGMKLIDFGIANQMOPDTTSIVKDSQVGTVNYMAPEAIRDMSSSRNSKIRT
KVSPRSDVMSLGCILYMTYGRTPFOHILNQVSKLHAIINPAHEIEFFPEISEKDLRDV
LKCLVRNPKERISIPELLTHPVQIQPHPGSQMARGATDEMKYVLQVLGLNSPNSI
LKTAKTLYERNCGEGQDSSSSKTFDKKRERK"
misc_feature 1567..2367
/note="S_TKc; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CDD:smart00220"

ORIGIN

Query Match 42.3%; Score 1634.6; DB 10; Length 2906;
Best Local Similarity 76.8%; Pred. No. 0;
Matches 2193; Conservative 0; Mismatches 559; Indels 103; Gaps 12;

QY 971 TGTAGAAATGGAATCCGAGGATTTAAAGTGGCAGAGAATTGACAATTGATCCATAATGAA 1030
|||
Db 69 TTTAGAAATGGAGGCTGAAGAGTTAATTGGCAGAGTGTGACGATTGATTCATCATGAG 128
|||
QY 1031 CAAAGTGAGAGACATTAATAATAAGTTTAAAAAATGAAGACCTTACTGATGAACCTT 1090
|||
Db 129 CAAAATGAGAGATATTAATAATAAG--ATAAATGAAGACTGTACTGATGAGCTAAGCTT 185
|||
QY 1091 GAATAAAATTTCTGCTGATACAGATAACTCGGGAACCTGTTAACCAAAATTATGATGAT 1150
|||
Db 186 GTCTAAATCTGTG-----CCGATCACACCGGAACCTGTTAACCAAAATTATGAGGGT 236
|||
QY 1151 GGCAAAACACCCAGAGGACTGTTGAGTTTGTGCTCAAAACTAGAGAAAAACAGTGTTC 1210
|||
Db 237 TGGGAACACCCAGAGAACTGTTGATTTCTTGTGTAAGACTGTACTGATGAGCTAAGCTT 296
|||
QY 1211 GCTAAGTATGCTCTTTTAAATAAATTGATTGTTGCTCGTTACAGTCAAGCAATTGAAGCGCT 1270
|||
Db 297 TCTAAATGACGATCTTTTAAATAAAGCTGATTGGTCGTATAGTCAAGCGGATTGAAGTACT 356
|||
QY 1271 TCCCCAGATAAATATGGCCAAAATGAGAGTTTTTGTAGAAATTCAGAGTGTGCTGA 1330
|||
Db 357 TCCTCCAGATAAATAACGGCCAGAAATGAGAGCTTTTGTCTGAATACAGTGAGACTTGCTGA 416
|||

Db	2508	TGGTGAAGGTCAAGATTCTTGGTCATCCAAAGACTTTTGACAAAAGAGAGAGAAAGATG	2567
Qy	3551	ATTTCAGATTATTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACTGTTATACTCT	3610
Db	2568	ATGCACAGCTACGTACAAACCAAGA-----ACACTAGATTGTTTCCCTCTGCCATACTCT	2621
Qy	3611	TGAATCCCTGTGGAATCTA-CATTTGAAGACAACATCACTCTGAAGTGT------	3660
Db	2622	TGAATCTCTGAGGAAATCTACCAGTTGGAACAACCTCACCTGGATTATCAGTTAAAA	2681
Qy	3661	--ATCAGCAAAAAAAATTCAGTGGAGATTATCTTTAAAGAAAACTGTAAAAATAGCAACC	3718
Db	2682	AAACAACAACAACAACAACCTTCAGTAGATTATCTCAAAAGAAAGCTGTAAAGTT---AACC	2738
Qy	3719	ACTTATGGCACTGTATATATTGTAGACTTGTTTTCT-CTGTTTTATGCTCTGTGTAATC	3777
Db	2739	ACTCATAGCACTGTGTATATTAAATTATAGAGTTGTGCTTTTCTTTTATGCTTTTCTGTA	2798
Qy	3778	TACTTGACATCATTTTACTCTTGGAAATAGTGGGTG	3812
Db	2799	AATCTGCTAATGTTTTTACGTTTGGAACACAGTGAATG	2833

RESULT 9
 AX775731
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 .
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1
 Matsuda, A. and Muramatsu, S.
 NF-kB activating gene
 Patent: WO 03048202-A 1 12-JUN-2003;
 Asahi Kasei Kabushiki Kaisha (JP)
 Location/Qualifiers
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES

CDS

ORIGIN

Query Match 42.2%; Score 1633; DB 6; Length 2860;
Best Local Similarity 76.8%;
Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 560; Indels 103; Gaps 12;
971 TGACAAATGGAATCCGACGATTTAAGTGGCAGAGAATTGACAAATGATTCCATAATGAA 1030

QY	2111	GGTTCCAGAGAGTAAC	CAGAAAA	CAGTGG	CAAGCTAAG	AAGAA	GTAG	AAGAA	GT	CAGAGT	GTATTA	AA	CA	2170
Db	1156	GATTGCAGAAA	GAAGAGGCC	CCATG	CAGTGG	CAGTCTAC	CCAG	AAAGCCG	GAGTGT	GTGT	TCCA	1215		
QY	2171	GAATCCTGCTGCAT	CTTCA	AAATCACTGG	CAGATTCCGG	AGTTAG	CCCG	AAAAAGTT	AA	TATAC	2230			
Db	1216	GAACCTGCTGCC	TTTGCA	CCCTTGCG	GCACGTTCC	AGATGTCA	CCCC	CGAAGG	CTGACAA	1275				
QY	2231	AGAGCAGAAA	CATACCA	CTTTTG	AGCAACCTG	CTTTTC	AGTTTCA	AAAA	CAGTCA	CCACC	2290			
Db	1276	AG-----									1287			
QY	2291	AATATCAACAT	CTAAATG	TTTGAC	CCAAAA	CTATT	TGTA	AGACAC	CAAGC	AGCAATAC	2350			
Db	1288	AATATCAGT	TCCTAA	ATGGCTTGAT	CCAA	AGTCTGCT	TGTG	AGACAC	CTTAG	TAGGAGCTC	1347			
QY	2351	CTTGATGATTAC	ATGAGCTG	TTTTAG	AACCTCC	AGTTGT	TAA	AGAA	TGACTTT	CCACCTGC	2410			
Db	1348	CTTGATGATTAC	ATGAAATG	TTTTA	AGACTCC	AGTTC	CAAG	ATGACTTT	CCACCTGC	1407				
QY	2411	TTGT	CAGTTGTCA	ACACCTTAT	TGGCCAA	CCTGCCTG	TTTCC	AGCAGCA	ACAGCAT	CAAAAT	2470			
Db	1408	CTGTCCAT	CATCAAC	ACCTTAC	AGCCAGCTT	GC	CTG	CCGCTCC	AGCAGCA	ACAGCAGGG	1467			
QY	2471	ACTTGCCACT	CCAC	TTCAAAATTTAC	AGGTTTTAG	CATCTTCTT	CAGCA	AAATGA	ATGCAT	2530				
Db	1468	ACTCAGCACT	CCTCTT	CAAGCTTGC	AGATTTT	CAGTTCTT	CATCA	TAAATGA	ATGCAT	1527				
QY	2531	TTCCGTTAA	AGGAAG	ATTTAT	TCCATATTA	AAAGCAGATAG	GAAGTGG	AGGTTCA	AGCAA	2590				
Db	1528	TTCAGTTAA	CGGAAG	ATTTAT	TCCATATTA	AAAGCAGATAG	GCAGTGG	AGGTTCC	AGTAA	1587				
QY	2591	GGTATTT	CAGGTGTTAA	ATGA	AAAAAGAAACAG	ATATATGCT	TATA	AAAAATATG	TGA	2650				
Db	1588	GGTGT	TTT	CAGGTA	TTGAATG	AGAAAAACAG	ATAAACGCTAT	CAAAATATG	TGA	1647				
QY	2651	AGAAGCAG	ATAACCA	AACTCTT	GTATAGTTAC	CGGAACG	AAAAATAGCTT	ATTGA	ATAAACT	2710				
Db	1648	AGAGCCG	ATAGCCAA	AACTATT	TGAGAGCTAC	CGCAACG	AGATAGCG	TTTTTTGA	CAAACT	1707				
QY	2711	ACAACAAC	CACAGTGA	TAAGATCAT	CCGACTTTAT	GATTA	TGAAATCA	CGGAC	CAGTACAT	2770				
Db	1708	ACAGCAAC	CACAGTGA	TAAGATCAT	CCGCTCTAT	GATTA	TGAAATCA	CGGAC	CAGTACAT	1767				
QY	2771	CTACATGG	TAAATG	GAGTGTG	AAATATTG	ATCTTA	ATAGTTGG	CTTAAAA	AGAAAAATC	2830				
Db	1768	CTACATGG	TAAATG	GGAATGTG	GAACA	ATTGAC	CTAAATAGTTGG	CTTAAAA	AGAAAAATC	1827				
QY	2831	CATTGATCC	ATGGSA	ACGCAAG	ATTACTG	SAAAAAATATG	TTAGAGG	CAGTTCA	CACAAT	2890				
Db	1828	CATCAAT	CCATGG	SAACGCAAG	AGCTACTG	SAAAAAACATG	TTGGAGG	CAGTAC	CATAAT	1887				
QY	2891	CCATCAAC	ATGSCAT	TGTTTCA	CGTGTCTTAA	ACCAGCTAA	CTTTCTG	ATAGTTG	ATGG	2950				
Db	1888	CCATCAG	CATG	GTATTTG	TTCATAGT	ATCTGA	AGCTGTAA	CTTTGTG	ATGTGG	1947				
QY	2951	AATGCTAA	AGCTAAT	TGATTTTGG	GATTGCA	AAACCAAT	TGCAAC	CAGATA	CAACAGTGT	3010				
Db	1948	AATGCTAA	AGCTAAT	TGATTTTGG	GATTGCA	AAACCAAT	TGCAAC	CAGATA	CAACAGTGT	2007				
QY	3011	TGTTAA	AGATTCTC	AGGTTGG	CACAGTTAA	TATATGCC	ACCAGAA	GCAATCAA	AGATAT	3070				
Db	2008	TGTTAA	AGATTCTC	AGGTTGG	CACAGTTAA	TATATGCC	ACCAGAA	GCAATCAA	AGATAT	2067				
QY	3071	GTCTT	CC	CAGAGAA	ATGGGAA	ATCTA	AGTCAA	AGATAAG	CCCCAAA	AGTGTG	3130			
Db	2068	GTCTT	CTTCA	AGAAAA	ATCGAAA	ATCAGG	ACCAAG	GTAA	GTCC	CAGAA	2127			
QY	3131	GTCTT	AGGATGT	ATTTTGT	ACTATAT	GACTTAC	GGGAAAA	ACCA	ATTT	CAGCAG	3190			
Db	2128	GTCTT	GGGGTGC	ATTTTGT	ACTACAT	GACTTAT	GGGAGG	ACGCC	ATTT	CAGCAG	2187			

QY	3191	TAATCAGATTCTCTAAATTAATCATGCCATAATTGATCTCTAATCATGAAATGAATTTCCCGA	3250
DB	2188	CAATCAGGTCTCTAAACTGACGCCATAATCAACCCGTCTCATGAGATTGAATTTCCCGA	2247
QY	3251	TATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTTGTTTAAAGAGGACCCCAAACA	3310
DB	2248	GATTCGGAAAAAGATCTTCGAGACGTGTAAAGTGCTGTTTAGTGAGGAACCCFAAGA	2307
QY	3311	GAGGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAACTCATCCAGT	3370
DB	2308	GAGGATATCTATCCCTGAGCTTCTCACACATCCGATGTTCAAATTCAGCCCCATCCAGG	2367
QY	3371	TAACCAAATGGCCAAAGGAACCACTGAAGAAATGAAATATGTTCTGGGCCAACCTGTTGG	3430
DB	2368	CAGCCAAATGGCTAGGGAGCCACTGATGAAATGAAATATGTTGTTGGGTCAACTGTTGG	2427
QY	3431	TCTGAATTCCTCACTCCATTTTGAAGCTGCTAAAACTTTATATGAACACTATAGTGG	3490
DB	2428	TCTGAATTCCTCACTCCATCTTGAAAACTGAAAAAATTTGTATGAACGTTATAATTG	2487
QY	3491	TGGTGAAGTCATAATTCCTTCATCCTCCAAGACTTTTGAAAAAAGGGGAAAAAATG	3550
DB	2488	TGGTGAAGTCAAGATTCCTTCGTCATCCAAGACTTTTGACAAAAGAGAGAAAGAAATG	2547
QY	3551	ATTTGCAGTTATTCGTAATGTCAGATAGGAGGTATAAAAATATATTGGACTGTTATACTCT	3610
DB	2548	ATGCACAGCTACGTACAAACCAAGA-----ACACTAGATTGTTTCTCTGCCATACTCT	2601
QY	3611	TGAATCCCTGTGGAAATCTA-CATTTGAAGACAACTCACTCTGAAGTGT-----	3660
DB	2602	TGAATCTCTGAGGAAATCTACCAGTTGGAAACAACTCACCTGGATTATCAGTTAAAA	2661
QY	3661	--ATCAGCAAAAAAATTCAGTGAGATTATCTTTTAAAGAAAACTGTAAAAATAGCAACC	3718
DB	2662	AAACAAACAAACAAAACCTTCAGTAGATTATCCTCAAAAGGAAGCTGTAAAGTT---AACCC	2718
QY	3719	ACTTATGGCACTGTATATATTGTAGACTGTTTTCT-CTGTTTTATGCTCTTGTGTAATC	3777
DB	2719	ACTCATAGCACTGTGTATATTAAATTTATAGAGTTGTCTTTTCTTTTATGCTTTTCTGTA	2778
QY	3778	TACTTGACATCATTTTACTCTTTGGAATAGTGGGTG	3812
DB	2779	AATCTGCTAATGTTTTACGTTTTAGACAGTGAATG	2813

RESULT 10	MUSESKK	2924 bp	linear	ROD 27-APR-1993
MUSESKK	LOCUS	2924 bp	mRNA	
DEFINITION	Mouse esk kinase mRNA, complete cds.			
ACCESSION	M86377			
VERSION	M86377.1 GI:193109			
KEYWORDS	esk kinase.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 2924)			
AUTHORS	Douville,E.M., Afar,D.E., Howell,B.W., Letwin,K., Tannock,L., Ben-David,Y., Pawson,T. and Bell,J.C.			
TITLE	Multiple cDNAs encoding the esk kinase predict transmembrane and intracellular enzyme isoforms			
JOURNAL	Mol. Cell. Biol. 12 (6), 2681-2689 (1992)			
MEDLINE	92269841			
PUBMED	1375325			
COMMENT	Original source text: Mus musculus (strain C3H) male cDNA to mRNA.			
FEATURES	Location/Qualifiers			
source	1..2924			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/strain="C3H"			
	/db_xref="taxon:10090"			
	/sex="male"			
	/cell_line="P19 EC"			

CDS		/cell_type="stem cell" 38..2608 /note="putative" /codon_start=1 /product="esk_kinase" /protein_id="AAA37578.1" /db_xref="GI:193110" /translation="MEAEELIGSSVTIDSIMSKWRDINKKINEDCTDELSLSKICADH TETVQIMRVGNTPENWLNFLKLEKNSSPLNDDLLNKLIGRYSQAIIEVLPDPKVGQN ESFARIQVRLAELKAIQEPDDARDYQFMARENCKKFAFVHVSFAQFELSQGNLKKSEQ LLHKAVETGAVPLQMLETAMRNHLQKKQLLPEEDKKSVSASTVLSAQEPFSSSLGNV QNRSTSCESRGAGAAVLYGENLPQDAEVRHQNPFFKQTHAAKRSQPFGRVPVNLN SPDFVKTDSAVTQLTIRLALSSVPLPYVTCUHLHQLLALAGLAKSGSPDRDAILPG SRPRGSDSYELRGLKPIOTIYLDLSVNEKSEHMSDLIALKSKTDSLSLTLEETKP EIAERPMQWQSTRKPECVFQNPAPAFAPLRHVDVTPKADKESPPISVPKWLDPKSAC ETPSSSLDDYMKCFKTPVVKNDFFPPACPSSTPYSQLARLQOQQOGLSTPLQSLQIS GSSSINECISVNGRIYSILKQIGSGSSKVFQVNLKKQINAIKYNLEDDADSQTIES YRNEIAFLNKLQOHSKDIIRLYDYEITEQYIYVMECGNIDLNLWKKKSINPWERK SYKNMLEAVHIIHQHGIVHSDLPANFVVDGMLKLIDFGIANQMPDTSIVKDSQ VGTVNYMAPEAIRDMSSSRNSKIRTKVSPRSVMSLGCILYVMTYGRTPFQHIINQV SKLHAIINPAHEIEFPEISEKDLRDVLKCLVRNPKERISIPPELLTHPYVQIQPHPGS QMARGATDEMKYVLGQLVGLNSPNSILKTAKTLYERYNCGEGQDSSSSSKTFDKKRERK "	
ORIGIN		Query Match Best Local Similarity 40.4%; Score 1562.4; DB 10; Length 2924; Matches 2194; Conservative 0; Mismatches 561; Indels 175; Gaps 12;	
QY	971	TGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAATTGATTCCTAATGAA	1030
Db	31	TTTAGAAATGGAGGCTGAAGAGTTAATGGCAGCAGTGTGACGATTGATTCCTCATGAG	90
QY	1031	CAAAAGTGAGAGACATTAATAAATAAGTTTAAAAATGAAGACCTTACTGATGAACTAAGCTT	1090
Db	91	CAAAATGAGAGATATTAATAAATAG---ATAAATGAAGACTGTACTGATGAGCTAAGCTT	147
QY	1091	GAATAAAATTTCTGCTGATACTACAGATAAATCTGGGAACTGTGTAAACCAATTTATGATGAT	1150
Db	148	GTCTAAAATCTGTG-----CCGATCACACCGAAACTGTGTAAACCAATTTATGAGGGT	198
QY	1151	GGCAAAACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCC	1210
Db	199	TGGGAAACACCCAGAGAACTGGTTGAATTTCTTGTGTAACCTAGAGAAAAACAGTCAACC	258
QY	1211	GCTAAGTGATGCTCTTTTAAATAAATTTGATGGTGGTTACAGTCAAGCAATTTGAAGCGCT	1270
Db	259	TCTAAATGACGATCTTTTAAATAAGCTGATTTGGTCCGTATAGTCAAGCGATTGAAGTACT	318
QY	1271	TCCCCAGATAAATATGGCCAAAATGAGAGTTTGTCTAGAATTCAAGTGAGATTTGCTGA	1330
Db	319	TCCTCCAGATAAATACGGCCAGAATGAGAGCTTTGCTCGAATACAACTAGAGCTTGTCTGA	378
QY	1331	ATTAAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTTCAAAATGGCCAGAGCAAA	1390
Db	379	ACTAAAGCTATTCAAGAGGCTGATGATGCCCGTGACTACTTCCAGATGGCCAGGAAAAA	438
QY	1391	CTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAAGTGTCAAGGTAA	1450
Db	439	CTGCAAGAAATTTGCTTTTGTGCACGTATCTTTTGCACAGTTTGAAGTGTCTCAAGGCCA	498
QY	1451	TGTCAAAAAAGTAACAACCTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGTA	1510
Db	499	TCTTAAAAAAGTGAGCAGCTTCTTCAATAAAGCTGTAGAGACTGGGGCGGTGCCGTGCA	558
QY	1511	AATGCTGGAATTTGCCCTGGCGAATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGGA	1570
Db	559	GATGCTGAGACGGCCATGCGTAACTTACACCTCCAGAAAAAGCAGCTGCTTCCGGAGGA	618
QY	1571	GGAAAAGAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATATTTCCGGTTC	1630
Db	619	GGACAGAAGAGTGTTGTACGATCGACAGTACTAAGTGCCCAAGAGCGCGTTCTCCAGCTC	678

QY	1631	ACTTGGGCATTTACAGAAATAGGAACAACAGTGTGTGATTTCCAGAGGACAGACTACTAAAGC	1690
Db	679	ACTTGGAAATGTACAGAAATAGGAGCATCAGCTGTGAGTCCAGAGGACAGGCTGGGGCAGC	738
QY	1691	CAGGTTTTTATATGGAGAGAACATGCCACCACCAAGATGCAGAAATAGGTTACCGGAATTC	1750
Db	739	CAGGGTTTTATATGGAGAGAACTTGCTCCACAAGATGCCAAGTCCGAAGTCAAAACCC	798
QY	1751	ATTGAGACAAACTPAACAAAACTAAACAGTCAATGCCCATTTTGGAAAGAGTCCCAGTTAACCT	1810
Db	799	CTTCAAGCAGACTCACGCAGCTAAACGGTCAATGCCCTTTTGGAAAGAGTCCCAGTCAATCT	858
QY	1811	TCTAAATAGCCCCAGATTGTGATGTGAAGACAGATGATTCACTGTGACCTTGTGTTTATGAA	1870
Db	859	TCTAAACAGCCCCAGATTCTATGTGAAGACAGATAGCTCAGCTGTGACACAGTTAACAAAC	918
QY	1871	AAGACAAACCTCTAGATCAGAAATG-----	1894
Db	919	AAGCTAGCCTTAAGCTCTGTACCTTGGCGTACGTAAACCTGCTCCTGCACATTACAGCT	978
QY	1895	-----CCGAGATTGGTTGTGCTGCTGG	1915
Db	979	GCTGGCGCTCGCAGGCTTGGCAAAGGGGTGAGGACCAGACCGGAGCGGATTCTGCCCGG	1038
QY	1916	ATCTAAACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAG	1975
Db	1039	CTCAGACCAACGTCGAGTGATTCCTATGAAGTGAAGGTTTAAAGCCCATTCAAACTAT	1098
QY	1976	TCATTTCAAGGAACCTCTGCTGTCAGATGAAAAGAGTTCTGAACITATTATTACTGATTC	2035
Db	1099	CTATTTGAAGACTCTTTGGTGTCCAATGAAAAGAGTTCTGAAC---TTATGCTGATTT	1155
QY	2036	AATAACCTTGAGAAATAAAACGGAAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGA	2095
Db	1156	AATAGCCTTGAAGAGTAAACACAGATTCAAG---TCTAACAAATTTGGAAGAACTA----	1208
QY	2096	GTATCAAGAACCCAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGCTAAGAGAAAGTC	2155
Db	1209	-----AGCCAGAGATTGCAGAAAGAGGCCCCATGCAGTGGCAGTCTACCAGAAAGCC	1260
QY	2156	AGAGTGATTAACCCAGAACTCTGCTGCATCTTCAAACTACTGGCAGATTCGGGAGTTAGC	2215
Db	1261	CGAGTGTGTGTTCCAGAACCTGCTGCTTTGCAACCCCTGCGGACCGTTCAGATGTAC	1320
QY	2216	CCGAAAAGTTAATACAGAGCAGAAACATACCACTTTTGAGCAACCTGCTCTTTCAGTTTC	2275
Db	1321	CCCGAAGGCTGA-----	1332
QY	2276	AAACAGTCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTAAGAC	2335
Db	1333	CAAAGAGTCAACCAATATCAGTTCCTAAATGGCTTGATCCAAAGTCTGCTTGTGAGAC	1392
QY	2336	ACCAAGCAGCAATACCTTGGATGATTACATGAGCTGTTTGAAGCTCCAGTTGTAAGAA	2395
Db	1393	ACCTAGTAGCAGCTCCTTGGATGATTACATGAATGTTTAAAGACTCCAGTTGTAAGAA	1452
QY	2396	TGACTTTCCACCTGCTTGTGAGTGTCAACACCTTATGGCCAAACCTGCTGTTTCCAGCA	2455
Db	1453	TGACTTTCCACCTGCTGCTCATCATCAACACCTTACAGCCAGCTTGCCCGCTCCAGCA	1512
QY	2456	GCAACAGCATCAATACCTTGGCACTCCACTTCAAAATTTACAGGTTTATAGCATCTTCTTC	2515
Db	1513	GCAACAGCAGGAGGACTCAGCACTCTCTTCAAGCTTGCAGATTTCAAGTCTTCTCATC	1572
QY	2516	AGCAATGAATGCATTTTCGGTTAAAGGAAGAAATTTTCCATATTAAAGCAGATAGGAAG	2575
Db	1573	AATAATGAATGCATTTTCAGTTAAACGGAAGAAATTTTCCATATTAAAGCAGATAGGCAG	1632
QY	2576	TGAGGTTCAAGCAAGGATTTTTCAGGTTGTTAAATGAAAAGAAACAGATATATGCTATAA	2635
Db	1633	TGAGGTTCCAGTAAGGTTTTCAGGTTATTTGAATGAGAAAAACAGATAAACGCTATCAA	1692
QY	2636	ATATGTGAACCTTAGAAGAGCAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGC	2695

Db 1693 ATATGTGAACCTAGAAGACCGCGATAGCAAACTATTGAGAGCTACCGCAACGAGATAGC 1752
QY 2696 TTATTTGAATAAATAACAAACACACAGTGAATAAGATCATCCGACTTTTATGATTATGAAAT 2755
Db 1753 GTTTTGAACAAACTACAGCAACACACAGTGAATAAGATCATCCGCTCTATGATTATGAAAT 1812
QY 2756 CACGGACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTTAATAGTTGGCT 2815
Db 1813 CACCGAGCAGTACATCTACATGGTAATGGAAATGTGGAACATTTGACCTAAATAGTTGGCT 1872
QY 2816 TAAAAAGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGA 2875
Db 1873 TAAAAAGAAAAATCCATCAATCCATGGGAACGCAAGAGTACTGGAAAAACATGTTGGA 1932
QY 2876 GGCAGTTTCACACAATCCATCAACATGGCATTTGTTCAAGTGTATCTTAAACAGCTAACTT 2935
Db 1933 GGCAGTACACATAATCCATCAGCATGGTATTGTTTCATAGTGTCTGAAGCCTGCTAACTT 1992
QY 2936 TCTGATAGTTGATGGAATGCTAAAGCTAATTGATTTGGGATTTGCAAAACCAATGCAACC 2995
Db 1993 TGTGATAGTGGATGGAATGCTAAAGCTAATTGATTTGGGATTTGCAAAACCAATGCAACC 2052
QY 2996 AGATACACAAGTGTGTTAAAGATTCTCAGGTTGGCACAGTTAATTATATGCCCACAGA 3055
Db 2053 AGACACAACAAGCATTTGTTAAAGATTCTCAGGTTGGCACAGTTAATTATATGCCCCAGA 2112
QY 3056 AGCAATCAAAGATATGCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCC 3115
Db 2113 AGCAATCAGAGACATGCTTCTTCAAGAGAAAAATTCGAAAAATCAGGACCAAGGTAAGTCC 2172
QY 3116 CAAAAGTGATGTTGGTCCCTTAGGATGATTTTGTACTATATGACTTACGGGAAACACC 3175
Db 2173 CAGAAGTGATGCTGGTCCCTTGGGGTGCAATTTGTACTACATGACTTATGGGAGGACGCC 2232
QY 3176 ATTTCCAGCAGATAAATTAATCAGATTTCTAAATTACATGCCATAATTGATCCTAATCATGA 3235
Db 2233 ATTTCCAGCAGATCATCAATCAGTCTCTAACTGCAGCCATAATCAACCTGCTCATGA 2292
QY 3236 AATTGAATTTCCGATATTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAA 3295
Db 2293 GATTGAATTTCCGAGATTTCCGAAAAAGATCTTCGAGACGTGTTAAAGTGTGTTTAGT 2352
QY 3296 AAGGGACCCAAAACAGAGGATATCCATTCTCTGAGTCTCTGGCTCATCCATATGTTCAAT 3355
Db 2353 GAGGAACCTTAAGAGAGGATATCTATCCCTGAGTCTCTCACACATCCGTATGTTCAAT 2412
QY 3356 TCAAACTCATCCAGTTAACCAATGGCCAAAGGGAACCACTGAAGAAATGAAATATGTTCT 3415
Db 2413 TCAGCCCATCCAGGCAGCCAAATGGCTAGGGAGCCACTGTATGAAATGAAATATGTTT 2472
QY 3416 GGGCCCACTTGTGGTCTGAAATCTCCTAACTCCATTTTGAAGCTGCTAAAACTTTTATA 3475
Db 2473 GGGTCAACTTGTGGTCTGAAATCTCCTAACTCCATCTTGAATACTGAAAACTTTTGT 2532
QY 3476 TGAACACTATAGTGGTGTGAAAGTCAATAATCTTCTCATCTCCTCAAGACTTTTGAACAAA 3535
Db 2533 TGAACCTTATAATTGTGTGAAGGTCAAGATCTTCTGTCATCCAGACTTTTGACAAAA 2592
QY 3536 AAGGGGAAAAAATGATTTGCGATTATTTCGTAATGTCAGATAGGAGGTATAAAATATATT 3595
Db 2593 GAGAGAAAGAAAGTGATGCACAGCTACGTACAAACCAAGA-----ACACTAGATTGTTT 2646
QY 3596 GGAAGTGTATCTCTTGAATCCCTGTGGAATCTA-CATTTGAAGACAAACATCACTCTGA 3654
Db 2647 CCTCTGCCATACTCTTGAATCTCTGAGGAATCTACCAAGTTGGAACCAACCTCACCTGA 2706
QY 3655 AGTGGT-----ATCAGCAAAAAAATTCAGTGAGATTATCTTTTAAAGAAACT 3703
Db 2707 TTTTATCAGTTAAAAAACAACAAACAAACTTCAGTAGATTATCTCTCAAAAGGAAGCT 2766
QY 3704 GTAAAAATAGCAACCACTTATGGCACTGTATATATTGTAGACTTGTCTTCT-CTGTTTATA 3762

Db 2767 GTAAAGTT---AACCACTCATAGCACTGTGTATATTAAATTATAGAGTTGTGCTTTTCTT 2823
QY 3763 TGCTCTTTGTGAATCTACTTGCATCATTTTACTCTTGAATAGTGGTG 3812
Db 2824 TTATGCTTTTCTGTAATCTGCTAATGTTTACGTTTGAACAGTGAATG 2873
RESULT 11
HSPYT
LOCUS HSPYT 946 bp mRNA linear PRI 30-JUN-1993
DEFINITION H.sapiens mRNA for phosphotyrosine picked threonine kinase (PYT).
ACCESSION X70500 S53714
VERSION X70500.1 GI:312815
KEYWORDS phosphotyrosine picked threonine kinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 946)
AUTHORS Lindberg,R.A., Fischer,W.H. and Hunter,T.
TITLE Characterization of a human protein threonine kinase isolated by screening an expression library with antibodies to phosphotyrosine
JOURNAL Oncogene 8 (2), 351-359 (1993)
MEDLINE 93149596
PUBMED 7678926
FEATURES
Location/Qualifiers
1..946
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
31..>946
/codon_start=1
/product="phosphotyrosine picked threonine kinase (pyt)"
/protein_id="CAA49912.1"
/db_xref="GI:312815"
/db_xref="GOA:Q15272"
/db_xref="SPTREMBL:Q15272"
/translation="MCMRQLQVLASSANECISVKGRIYSILKQIGSGGSKVQVL
NEKQIYAIKYVNLAEADNOTLDSYRNEIAYLNKLOQHSKIRLYDEITDOYLYMV
MECGNIDLNSWLKKKSIDPWERKSYWNMLEAVHTIHOHIVHSDLKPANFLIVDGM
LKLIDFGIANOMQDPTTSVKDSQVGTVNYMPPEAIKMSRSSRENGSKSKISPKSDV
WSLGCILYMTYKTPFPQIINQISKLHAIIDPNHEIEFDIPEKDLQDVVXCKLKR
PKQRISIPPELLAHYPVYQIOTHPVQNMAGI"

ORIGIN

Query Match 23.2%; Score 895.4; DB 9; Length 946;
Best Local Similarity 99.9%; Pred. No. 1.3e-177;
Matches 896; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2487 CAAATTTACAGGTTTTCAGCATCTTCTTCAGCAATGAATGCAATTCGGTTAAAGGAAGA 2546
Db 43 CAAATTTACAGGTTTTCAGCATCTTCTTCAGCAATGAATGCAATTCGGTTAAAGGAAGA 102
QY 2547 ATTTATTCATATTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGTATTTCAGGTGTTA 2606
Db 103 ATTTATTCATATTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGTATTTCAGGTGTTA 162
QY 2607 AATGAAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGAACAGATAACCAA 2666
Db 163 AATGAAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGAACAGATAACCAA 222
QY 2667 ACTCTGTAGATTACCGGAACGAAATAGCTTATTGTAATAACTACAAACACAGTGAT 2726
Db 223 ACTCTGTAGATTACCGGAACGAAATAGCTTATTGTAATAACTACAAACACAGTGAT 282
QY 2727 AAGATCATCCGACTTTATGATTATGAAATCAGGACCAGTACATCTACATGGTAATGGAG 2786
Db 283 AAGATCATCCGACTTTATGATTATGAAATCAGGACCAGTACATCTACATGGTAATGGAG 342
QY 2787 TGTGAAATATTGATCTTAATAGTTGGCTTAAAAAGAAAAATCCATTGATCCATGGGAA 2846
Db 343 TGTGAAATATTGATCTTAATAGTTGGCTTAAAAAGAAAAATCCATTGATCCATGGGAA 402

QY 2847 CGAAGAGTTACTGTGAAATAATATGTTAGAGGCAGTTTCACAAATCCATCAACATGGCATT 2906
Db |||||
QY 403 CGAAGAGTTACTGTGAAATAATATGTTAGAGGCAGTTTCACAAATCCATCAACATGGCATT 462
Db |||||
QY 2907 GTTCACAGTGTCTTAAACACGCTAACTTTCTGATAGTTGATGAATGCTTAAAGCTAATT 2966
Db |||||
QY 463 GTTCACAGTGTCTTAAACACGCTAACTTTCTGATAGTTGATGAATGCTTAAAGCTAATT 522
Db |||||
QY 2967 GATTTGGGATTGCAAAACCAATGCAACACGATACAAACAGTGTGTTAAAGATTCTCAG 3026
Db |||||
QY 523 GATTTGGGATTGCAAAACCAATGCAACACGATACAAACAGTGTGTTAAAGATTCTCAG 582
Db |||||
QY 3027 GTTGGCACAGTTAATTATATGTCACAGAAAGCAATCAAGATATGTTCTCTCCAGAGAG 3086
Db |||||
QY 583 GTTGGCACAGTTAATTATATGTCACAGAAAGCAATCAAGATATGTTCTCTCCAGAGAG 642
Db |||||
QY 3087 AATGGAAATCTAAGTCAAAGATAAGCCCAAAAGTGATGTTGGTCTTTAGGATGTAAT 3146
Db |||||
QY 643 AATGGAAATCTAAGTCAAAGATAAGCCCAAAAGTGATGTTGGTCTTTAGGATGTAAT 702
Db |||||
QY 3147 TTGTACTATATGACTTACGGGAAACACCATTTTCAGCAGATAATTAATCAGATTCTAAA 3206
Db |||||
QY 703 TTGTACTATATGACTTACGGGAAACACCATTTTCAGCAGATAATTAATCAGATTCTAAA 762
Db |||||
QY 3207 TTACATGCCATATGATCCTTAATCATGAATTTGAATTTCCGATATTTCCAGAGAAAGAT 3266
Db |||||
QY 763 TTACATGCCATATGATCCTTAATCATGAATTTGAATTTCCGATATTTCCAGAGAAAGAT 822
Db |||||
QY 3267 CTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCCAACAGAGGATATCCATTCCT 3326
Db |||||
QY 823 CTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCCAACAGAGGATATCCATTCCT 882
Db |||||
QY 3327 GAGCTCCTGGCTCATCCATATGTTCAAATTTCAAATTCATCCAGTTAACCAATGGCC 3383
Db |||||
QY 883 GAGCTCCTGGCTCATCCATATGTTCAAATTTCAAATTCATCCAGTTAACCAATGGCC 939
Db |||||

RESULT 12
AC112218 144319 bp DNA linear PRI 24-MAY-2002
LOCUS Homo sapiens chromosome 3 clone CTD-2036B16, complete sequence.
DEFINITION AC112218
ACCESSION AC112218
VERSION AC112218.2 GI:21166214
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 144319)
AUTHORS Kaul,R.K.; Olson,M.V.; Zhou,Y.; James,R.A.; Rouse,G.; Wu,Z.;
Saenphimmachak,C.; Phelps,K.A.; Buckley,D.; Raymond,C. and
Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 144319)
AUTHORS Kaul,R.K.; Olson,M.V.; Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 144319)
AUTHORS Kaul,R.K.; Olson,M.V.; Zhou,Y.; James,R.A.; Rouse,G.; Wu,Z.;
Saenphimmachak,C.; Phelps,K.A.; Buckley,D.; Raymond,C. and
Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On May 24, 2002 this sequence version replaced gi:18767517.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgchgs@u.washington.edu
----- Project Information

Center project name: chr-3
Center clone name: CTD-2036B16 (bc0724)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator ET; 57% of reads
Chemistry: Dye-terminator Big Dye; 43% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 144239 bases at least Q40
Consensus quality: 144315 bases at least Q30
Consensus quality: 144319 bases at least Q20
Insert size: 144319; sum-of-contigs
Quality coverage: 10.3x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': CTD-2200P7 (UWGC:bc0739)
3': Mapping in progress

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.

EcoRI				BglII				HindIII			
SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
3200	3234	4451	4381	7510	7532						
-----	-----	-----	-----	-----	-----						
6633	6681	2067	2050	19977	19743						
-----	-----	-----	-----	-----	-----						
75	<800	3747	3776	3009	3037						
-----	-----	-----	-----	-----	-----						
13226	12873	93	<800	8698	8821						
-----	-----	-----	-----	-----	-----						
12985	12873	8674	8581	378	<800						
-----	-----	-----	-----	-----	-----						
2918	2905	1940	2050	4338	4345						
-----	-----	-----	-----	-----	-----						
17848	18128	9569	9330	8944	8821						
-----	-----	-----	-----	-----	-----						
2843	2905	755	746	1891	1899						
-----	-----	-----	-----	-----	-----						
10469	10380	3577	3561	2502	2484						
-----	-----	-----	-----	-----	-----						
1195	1188	1018	1038	408	<800						
-----	-----	-----	-----	-----	-----						
1938	1918	2555	2620	1361	1345						

220	<800	2284	2324	8221	8317
-----	-----	-----	-----	-----	-----
7669	7740	2355	2324	751	810
-----	-----	-----	-----	-----	-----
2870	2917	1784	1778	661	<800
-----	-----	-----	-----	-----	-----
312	<800	2786	2825	963	967
-----	-----	-----	-----	-----	-----
3724	3732	2583	2637	3985	3949
-----	-----	-----	-----	-----	-----
5021	4915	2055	2005	3618	3548
-----	-----	-----	-----	-----	-----
984	968	25611	26061	3482	3548
-----	-----	-----	-----	-----	-----
85	<800	1055	1055	4746	4669
-----	-----	-----	-----	-----	-----
7133	7070	2196	2211	488	<800
-----	-----	-----	-----	-----	-----
251	<800	1463	1414	1050	1047
-----	-----	-----	-----	-----	-----
3194	3240	359	<800	6492	6360
-----	-----	-----	-----	-----	-----
9434	9351	75	<800	8735	8934
-----	-----	-----	-----	-----	-----
4049	4031	2959	2825	786	810
-----	-----	-----	-----	-----	-----
2914	2917	6536	6511	5628	5568
-----	-----	-----	-----	-----	-----
1381	1344	714	<800	2417	2421
-----	-----	-----	-----	-----	-----
5262	5244	232	<800	8446	8317
-----	-----	-----	-----	-----	-----
		316	<800	2645	2635
-----	-----	-----	-----	-----	-----
		5299	5189	4199	4186
-----	-----	-----	-----	-----	-----
		19353	19040	1629	1568
-----	-----	-----	-----	-----	-----
		482	<800	1801	1937
-----	-----	-----	-----	-----	-----
		160	<800	2040	2009
-----	-----	-----	-----	-----	-----
		1803	1778	1786	1774
-----	-----	-----	-----	-----	-----
		8	<800	4387	4186
-----	-----	-----	-----	-----	-----
		1586	1518	13666	13356
-----	-----	-----	-----	-----	-----
		4087	4079		
-----	-----	-----	-----	-----	-----
FEATURES	Location/Qualifiers				
source	1. .178596				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="3"				
	/clone="RP11-82B23"				
	/clone_lib="RPC1 human BAC library 11"				
misc_feature	165805..165810				
	/note="Single subclone region"				
ORIGIN					
Query Match	22.3%; Score 863.6; DB 9; Length 178596;				
Best Local Similarity	98.6%; Pred. No. 4.9e-171;				
Matches 925; Conservative	0; Mismatches 4; Indels 9; Gaps 5;				
QY	4 ATTCCTTTT	TTTTTTT	TGAGATGGAGTTT	CACCTT	GTGTCAGGCTGGAGTCAA 63
Db	159373 ATGACTTTT	TTTTTTT	TGAGATGGAGTTT	CACCTT	GTGTCAGGCTGGAGTCAA 159432
QY	64 TGGCACAATCT	CAGCTT	ACTGCAACCT	CCGCTCCCGGGTT	CAAGCGATTCTCCTGCCTC 123
Db	159433 TGGCACAATCT	CAGCTT	ACTGCAACCT	CCGCTCCCGGGTT	CAAGCGATTCTCCTGCCTC 159492

QY	124	AGCCTCTCAAGTAGCTGGATTACAGGCATGTGCCACCACCCCTGGCTAACTAATTTCTT	183
Db	159493	AGCCTCTCAAGTAGCTGGATTACAGGCATGTGCCACCACCCCTGG-----CTAATTTCTT	159548
QY	184	TTCTATTAGTAGAGATGGGGTTTCAACCATGTTGGTCAGGCTGGTCTTGAACTCCTGACC	243
Db	159549	TTCTATTAGTAGAGATGGGGTTTCAACCATGTTGGTCAGGCTGGTCTTGAACTCCTGACC	159608
QY	244	TCAGGTGATCCACTTGGCTTGGCCCTCCCAAAGTCTAGGATTACAGCGGTGA--AACTGT	301
Db	159609	TCAGGTGATCCACTTGGCTTGGCCCTCCCAAAGTCTAGGATTACAGCGGTGAGCCACTGT	159668
QY	302	GCCTGGCTGATTC-TTTTTTTTGTGTTGGATTGTTTGAACAGGGTCTCCCTTGGTCGCC	360
Db	159669	GCCTGGCTGATTC-TTTTTTTTGTGTTGGATTGTTTGAACAGGGTCTCCCTTGGTCGCC	159728
QY	361	AGGCTGGAGTGCAGTGGTGGGATCTTGGCTCACTATAACCTCCACCTCTGGTTTCAAGT	420
Db	159729	AGGCTGGAGTGCAGTGGTGGGATCTTGGCTCACTATAACCTCCACCTCTGGTTTCAAGT	159788
QY	421	GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGACCAACACCCCG	480
Db	159789	GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGACCAACACCCCG	159848
QY	481	CTAATTTTGTATTTTATTAGAGACAGGGTTTCAACCATGTTGGCCAGGCTGTTCTCAA	540
Db	159849	CTAATTTTGTATTTTATTAGAGACAGGGTTTCAACCATGTTGGCCAGGCTGTTCTCAA	159908
QY	541	CTCCTGGACTCAAGGGATCCGCTGCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG	600
Db	159909	CTCCTGGACTCAAGGGATCCGCTGCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG	159968
QY	601	AGTCACCATGCCTGACCTTATAATTTCTTAAGTCA-TTTTTTCTGGTCCATTTCTTCTTA	659
Db	159969	AGTCACCATGCCTGACCTTATAATTTCTTAAGTCA-TTTTTTCTGGTCCATTTCTTCTTA	160028
QY	660	GGTCTCTCACAAATCTGCATTAGGCGGTACAATAATCTTAACTTCATGATTACAA	719
Db	160029	GGGT-CTCACAAATCTGCATTAGGCGGTACAATAATCTTAACTTCATGATTACAA	160087
QY	720	AAGGAAGATGAAGTGAATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACATCC	779
Db	160088	AAGGAAGATGAAGTGAATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACATCC	160147
QY	780	TGGATGATGATCCTAAATCCAGATACAGTAATAATGGGGTATGGGAAGGTAGAAATACAA	839
Db	160148	TGGATGATGATCCTAAATCCAGATACAGTAATAATGGGGTATGGGAAGGTAGAAATACAA	160207
QY	840	ATTTGGTTTAAATTAATTAATCTAAATATCTAATAATCTAATAAATTTTGGATGATG	899
Db	160208	ATTTGGTTTAAATTAATTAATCTAATAATCTAATAAATTTTGGATGATGATGATG	160267
QY	900	AATGTAAGACTGTACAGACTTCTCTAGAAAACAGTTTGG	937
Db	160268	AATGTAAGACTGTACAGACTTCTCTAGAAAACAGTTTGG	160305

RESULT 14
AF250290
LOCUS AF250290 2649 bp mRNA linear VRT 04-DEC-2001
DEFINITION Xenopus laevis Mps1/TTK mRNA, complete cds.
ACCESSION AF250290
VERSION AF250290.1 GI:13491130
KEYWORDS
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 2649)
AUTHORS Abrieu,A., Magnaghi-Jaulin,L., Kahana,J.A., Peter,M., Castro,A.,
Vigneron,S., Lorca,T., Cleveiland,D.W. and Labbe,J.C.

TITLE Mps1 is a kinetochore-associated kinase essential for the vertebrate mitotic checkpoint
JOURNAL Cell 106 (1), 83-93 (2001)
MEDLINE 21354313
PUBMED 11461704
REFERENCE 2 (bases 1 to 2649)
AUTHORS Magnaghi-Jaulin,L., Vigneron,S., Lorca,T. and Labbe,J.C.
TITLE X-Mps1/TTK is a novel Xenopus protein required for the spindle checkpoint
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2649)
AUTHORS Magnaghi-Jaulin,L., Vigneron,S., Lorca,T. and Labbe,J.C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) CRM, CNRS, 1919 route de Mende, Montpellier cedex 34293, France
FEATURES
source
Location/Qualifiers
1..2649
/organism="Xenopus laevis"
/mol_type="mrna"
/db_xref="taxon:8355"
1..2649
/codon_start=1
/product="Mps1/TTK"
/protein_id="AAK27843.1"
/db_xref="GI:13491131"
/translation="MDEDISERKLKIASILDRVSPKTYGTDNWDDELTFKSSA
DTTEHSGIFTHLVTKTPEWLNCLLKLENTGLPQIDPOLLNKLIDNYSOAVGALPAE
KHSNYESYAKILVHFAELKAIHDLDEAREQFOFARLNCKKFAFVHTAPAQFELSEGNF
KKCKOILRRGLECGAVPCMLNALKVQFKPOLISDEDKENLAVSSSHINQGTASF
QNLALGNPQRMKIESPEEYSVKTRFSYGEKLSRPEDFEDIGRKLNNMSAKTCPLGRV
PVQPATSPDTRKSDGSGSSVVKRPFSMSRPVPLATLSNPKVLGDDLHCSEDIKVP
SSNILPHEEQANEDSLDMKTPSSVILSVTESNLTVKRNEDLILGSTSTAVLNPNQENSK
PSESAPLLPNINARLSDTKPNNIEAEWKVKVPETPRQIFQPEMIFPAGKRKSTEPS
ANPGSRVSPAPSLSKDPVFCVCGTPVNSQEDYMNCFRTPVVTNGLPLAQMSTPY
NNRSDYQOPQTPVGQPCFISAPFASDDCIUVKGRAYAVLKQIGTGSSKVFQVMD
KKHLVAIKYVNLEEAQQTIYESYQNEISHLNKLQOHCXKIIIRLFDEYITDQHIYVMVE
CGNIDLNTWLRKKKTINPWRKSYWKWMLAEVTHIQHGIHVSDLKPANFLIVDGMK
LIDFGIANQIQPDVTSIVKDSQVGTINYPMPESIRDTTSYAENGKPRSKISPKGDVWS
LGCILYCMTYGKTPFQHTNQIAKLHSILDPGYEIEFPNIPKLDQDLVKCLVRNPK
ERISAEILLVHPYVQIQPHTQPDQVQKETTEEMKRILQGLIGLNSPNSISRAAKNLY
DQFNSGRSLDLSTLGLTTVSQNTRTTK"
ORIGIN
Query Match 20.5%; Score 793.8; DB 5; Length 2649;
Best Local Similarity 60.2%; Pred. No. 2.6e-156;
Matches 1575; Conservative 0; Mismatches 947; Indels 93; Gaps 12;
QY 978 ATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAATTGATTCCATAATGAACAAAGTG 1037
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 ATGGATGATGAAGATATAAGTGAACGGAACTGAAATTTGTTCTATCTTTGGTAGAGTT 60
QY 1038 AGAGACATTAATAATAAGTTTAAATAATGAAGACCTT---ACTGATGAACCTAAGCTTGAAT 1094
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 AAGAGTTTAAAGACCAAGTATGGAACAGATGATAATTGGACTGATGAATTAACCTTTCA 120
QY 1095 AAAATTTCTGCTGATACTACAGATAACTCGGGAACTGTTTAAACCAAAATTATGATGATGGCA 1154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 AAGTCTTCTGCTGACACTACAGAGCACTCTGGCATTTTACCCTA---CTTGGTGACCACA 177
QY 1155 AACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCGGCTA 1214
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
178 AAGACTCCTGAGGAATGGCTAAATTGCTTCTTAAATTTAGAAAAATACAGGCCTGCCTCAG 237
QY 1215 AGTGATGC---TCTTTTAAATAAATTGATTGGTTCGTTACAGTCAAGCAATTGAAGCGCTT 1271
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
238 ATTGACCCACAACCTTCTCAATAAGCTCATTGATAATTACAGTCAAGCTGTTGGAGCCCTG 297
QY 1272 CCCCCAGATAAATAATGGCCAAAATGAGAGTTTGTGCTAGAAATTCAAGTGAGATTTGCTGAA 1331
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 CCGCAGAAAAAACAGTCACAATGAGAGCTATGCTAAATTCCTGTGCACTTTGCGCAA 357
QY 1332 TTAAAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGGCCAGAGCAAAAC 1391
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 358 TTGAAAGCAATCCATGACCTAGATGAACACGGGAGCAGTTCAGTTTCCAAGACTAAAC 417
QY 1392 TGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACACTGTCAAGGTAAT 1451
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
418 TGCAAAAAGTTTGTGTCATACAGCTTTTGGCCAAATTTGAGCTGTCAAGAGGAAAT 477
QY 1452 GTCAAAAAAGTAACAACTTCTTCTCAAAAAGCTGTAGAACGTGGAGCAGTACCACCTAGAA 1511
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
478 TTTAAAAAGTGTAAAGCAGATTCTTCGAAGAGGTCTGGAATGTGGCGCTGTTCCCTCGGAA 537
QY 1512 ATGCTGGAAATTCCTCGCGGAAATTTAAACCTCCAAAAAAGCAGCTGCTTTCAGAGGAG 1571
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
538 ATGTTGAACCTTGGCTTTGAAGAATGTACAGTTTAAAGAGCCCTCAGCTGATATCTGACGAA 597
QY 1572 GAAAAAGAAATTTATCAGCATC-TACGTTATTAACTGCCCAAGAAATCATTTTCCGGTTC 1630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
598 GATAAAGAAAAATTTAGCAGTGTCTCAAGTCATATCAATCAAGGCACGGCTAGTTTTCAG 657
QY 1631 ACTTGGGCATTTACAGAANTAGGAACACACAGTTGTGATT-----CCAGAGGACAGACTACT 1685
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
658 AATTGGCACTGGGAAATCCACAACGAATGAAGATCGAAAGCCCAAGAAATATTCTGTGA 717
QY 1686 AAAGCCAGGTTTATATATGGAGAGAACATGCCACCAAGATGCAGAAAAATAGGTTACCGG 1745
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 AAAACAAGATTTTCATATGGAGAAAAAGTTAAGCTCCCCAGAAAGACTTTTGAAGATATAGGA 777
QY 1746 AATTCAATTGAGACAAAATAACAAAACATAACAGTCATGCCCAATTTGGAAGAGTCCCAGTT 1805
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
778 AGAAACCGTTTACT--TAACATGTCTGCTAAGACCTGTCTTTGGGCAGAGTGCCCTGTA 834
QY 1806 AACCTTCTAAATAGCCAGAT--TGATGTGTGAAGACAGATGATTCAGTTGTACCTTGT 1862
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
835 CAACCAGCAACATCTCCAGATACAAGAACAAAGGAAGAGTGTGCTGCGCAGTTCAGT 894
QY 1863 TTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTTGGTTGTGCTGGATCTAAA 1922
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
895 GTTGTGAAGAGACCATTCAGTTCCATGCGTGTGCTGTCTCTTGTACTTTTGTCAAATCCA 954
QY 1923 CCAAGTGGAAATGATCTCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATTTT 1982
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
955 AAATATTTGGGAGATGATCTGCATTGCTCAGAAAGACATTAAAGTACCAAGCAGCAACATC 1014
QY 1983 AAGGAACCTCTGGTGTCAAGATGAAAAGAGTTCTTGAACTTATTATTACTGATTCAATAACC 2042
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1015 CTACCACATGAAGAACAGGCTAATGAAGACAGCTTAGATATGAAGACACCGTCTCTGT 1074
QY 2043 CTGAAGAATAAACCGAATCAAGTCTTCTAGCTAAATTTAGAGAA----- 2087
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1075 ATTCTTAGTGTCAAGAAATCAACCTGACTGTCAAAAGAAATGAAGATTTAATTTTGGGA 1134
QY 2088 -----ACTAAAGAGTATCAAGAACCGAGGTTCAGAACCGAGGTTCCAGAGAGT 2123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1135 AGCAGCTCTACAGCTGTTCTCAATCTCAAGAAAAACAGTAAACCATCTGAATCAGCACCA 1194
QY 2124 AACCAGAAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGTGTATTAAACCAGAAATCCTGTGCA 2183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1195 CCACCTTTTACCCCAACATTAATGCAAGGCTCTCTGACACGACCAAGCCAAAACAATATTGAA 1254
QY 2184 TCTTCAAATCACTGGCAGATTCCGGAGTTAGCCCGAAAAGTT----- 2225
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1255 GCAGAAATGGAAGTGAAGGTTCCAGAGACACCCAGGCAATTTTCCAGCCAGAGATGATA 1314
QY 2226 AATACAGAGCAGAAACATACCCTTTTGTAGCAACCTGTCTTTTTCAGTTTCAAAACA---G 2282
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1315 TTTCCAGCGGGTAAAAGAAAAGAGTACTGAGCCTAGTGTCTAATCCTGGTTCCTCCGACGAGTG 1374
QY 2283 TCACCACCAATATCAACATCTAAATGGTTTGACCCCAAAATCTATTGTAAAGACACCAAGC 2342
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1375 TCTCCGCCAGCACCTCTCTTAGTAAATGTGACCTGTCTTTTGTGTGGAACACCCAGTC 1434
QY 2343 AGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTCAGTTGTAAAGAATGACTTT 2402
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1435 AATAAATCACAAAGAGGATTACATGAACCTGTTTAGGACTCCAGTTGTAAAAACGAACCTA 1494

QY 2403 CCACCTGCTTGTTCAGTTGTCACACACCTTATGCGCAACCTGCCTGTTTCCAGCAGCAACAG 2462
Db |||||
1495 GGACCAATGGCACAATGTCACACCCCTACAACAACCGTTTCAGATTATCAGCAACCCACAG 1554

QY 2463 CATCAATAACTTCCACTCCACTTCAAAATTTACAGGTTTATAGCATCTTCTTTCAGCAAT 2522
Db |||||
1555 ACCCCAGTTGGA-----CAACCGGACTGTTTCCAAATCTCTGCACCTTTTGCATCAAGT 1608

QY 2523 GAATGCATTCGGTTAAAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGAGGT 2582
Db |||||
1609 GACTGTATTGTAGTTAAAGGACGTGCATATGCAGTTTAAACAATAATGGAACAGGTGGA 1668

QY 2583 TCAAGCAAGGTATTTCAGGTGTTAAATGAAAGAAACAGATATATGCTATAAAATATGTG 2642
Db |||||
1669 TCTAGTAAGGTGTTTCAAGTAATGGATGATAAAAGCATTTATATGCCATAAAGTATGTG 1728

QY 2643 AACTTAGAAGAAGCAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATTG 2702
Db |||||
1729 AATCTCGAAGAAGCTGACCAGCAGACCATAGAAAGCTATCAGATGAAATTTCTCACTG 1788

QY 2703 AATAAATCAACAACACACTGATAAGATCATCCGACTTTATGATTATGAAATCACGGAC 2762
Db |||||
1789 AACAACTTCAGCAACACTGTGACAAGATCATTAGGCTCTTTGACTATGAGATTACTGAC 1848

QY 2763 CAGTACATCTACATGGTATGGAGTGTGGAATATTGATCTTAAATAGTTGGCTTAAAG 2822
Db |||||
1849 CAGCATATTTACATGGTAATGGAGTGTGGAATATAGATCTCACTACTTGGCTAAGGAAG 1908

QY 2823 AAAAAATCCATTGATCCATGGGAACGCAAGAGTACTGGAATAATATGTTAGAGGCAGTT 2882
Db |||||
1909 AAAAAACTATTAATCCTTGGGAACGGAAGAGCTACTGGAATAATATGTTGAAGCAGTC 1968

QY 2883 CACACAATCCATCAACATGGCATTTGTCACAGTATCTTAAACACAGCTAACTTCTGATA 2942
Db |||||
1969 CACACAATACATCAGCATGGGATTGTACACAGTGACTTGAAGCCAGCAAACTTCTTAAT 2028

QY 2943 GTTGATGGAATGCTAAAGCTAATGATTTTGGGATGCAACCAATGCAACCAAGTACA 3002
Db |||||
2029 GTTGACGGGATGCTCAAGCTAATAGACTTTGGCATAGCAAAACCAATCCAGCCAGATGTA 2088

QY 3003 ACAAGTGTGTTAAAGATTCTCAGGTGGCAGTTAAATTATATGCCACCAAGCAATC 3062
Db |||||
2089 ACCAGCATTTGTTAAAGATTCAACAGGTGGGACGATTAAATTACATGCCACCAAGATCTATC 2148

QY 3063 AAAGATATGTTCTTCCAGAGAGAAATGGAATCTAAGTCAAGATAAGCCCAAAAGT 3122
Db |||||
2149 AGAGATACCACCTCATATGCGGAAATGGAACCCAGATCAAGATAGCCCAAAAGGT 2208

QY 3123 GATGTTGGTCTTAGGATGTTATTTGTTACTATATGACTTACGGGAAACACCACTTTCAG 3182
Db |||||
2209 GATGCTGGTCTCTGGGTTGCATATTGTTACTGCATGACATATGGAAGACTCCTTTTCAA 2268

QY 3183 CAGATAAATTAATCAGATTCTTAAATTAATACATGCGCATAATTGATCCTAATCATGAAATGAA 3242
Db |||||
2269 CATATTACAAATCAGATTGCTAAACTTCAATTCATCTGGATCCTGGTTATGAAATAGAG 2328

QY 3243 TTTCCCGATATTCAGAGAAAGATCTTCAAGATGTTGTTAAAGTGTGTTTAAAGAGGAC 3302
Db |||||
2329 TTCCCTAATATACCAGAGAAAGACCTTCAAGATGTAAGGAATGTTTAGTACGAAT 2388

QY 3303 CCAAAACAGAGGATATCCATTCTTGAGCTCCTGGCTCATCCATATGTTCAAAT-----T 3356
Db |||||
2389 CCTAAGGAAGGATATCCATTGCTGAGCTCTTGTGCACCCATATGTTTCAGATACAGCCT 2448

QY 3357 CAAACTCATCCAGTTAAACAAATGGCCAAAGGAACCACTGAAGAAATGAAATATGTTCTG 3416
Db |||||
2449 CATACACAGCCAGATCAACAAGTGCAAAAGACAGACGACAGAGAAATGAAGCGTATATC 2508

QY 3417 GGCAACTTGTGTTGATCTGAAATCTCCTAATCTCCATTTTGAAGTGTCTAATACTTATAT 3476
Db |||||
2509 GGACAGCTTATTGGTCTGAATTTCTCCAAATTCATATCGCGAGCTGCAAAAAAATTTGTAT 2568

QY 3477 GAACACTATAGTGGTGAAGTCAATATCTTC 3511
Db |||||
2569 GATCAGTTTAAACAGTGGCAGAAGCCTGGATCTGTC 2603

RESULT 15
AC022415/c
LOCUS AC022415 180510 bp DNA linear PRI 22-DEC-2000
DEFINITION Homo sapiens chromosome 19 clone CTC-359D24, complete sequence.
AC022415
VERSION AC022415.5 GI:11968305
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 180510)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
2 (bases 1 to 180510)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 180510)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (22-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Dec 22, 2000 this sequence version replaced gi:7704995.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
STS Content:
SHGC-6150 G14148.
Location/Qualifiers
1. 180510
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-359D24"

ORIGIN
Query Match 10.8%; Score 418; DB 9; Length 180510;
Best Local Similarity 81.1%; Pred. No. 1.8e-77;
Matches 498; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 9 TTTTCTTTTCTTTTTCAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGCATGGCA 68
Db |||||
129549 TTATTTTATTTTTCAGACGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGCATGGTG 129490

QY 69 CAATCTCAGCTTACTGCAACCTCCGCTCCCGGGTTCAAGCATTCCTGCTCAGCT 128
Db |||||
129489 CAATCTCAGCTCACCGCAATCTCTGCTCCCGGGTTCAAGTGAATTCCTGCTCAGCT 129430

QY 129 CTCAAGTAGCTGGATTACAGGCATGTGCCACCACCCCTGGCTAACTATTTCTTTCTA 188
Db |||||
129429 CCAAGTAGCTGGATTACAGGCATGTGCCAACATGCCAGCTAAATGTGATTTT 129370

QY 189 TTTAGTAGAGATGGGTTTCAACATGTTGGTCAGGCTGGTCTTGAACCTCCTGACCTCAGG 248
Db |||||
129369 TTTAATAGAGACAGGTTTCTCCATGTTGGTCAGGCTGGTTCGAACTCCGACCTCAGG 129310

QY 249 TGATCCACTTGCCTTGGCTCCCAAGTGTAGGATTACAGCGTGAACCTGTGCTGCGC 308
Db |||||
129309 TGATCCGCCACCTTGGCTCCCAAGTGTGGATTATAGCATGAGCCACTGCGCCTG 129250

QY 309 TGATCTTTTCTTTGTTGGATTTTTGAACAGGGTCTCCCTTGGTCGCCAGGCTGGA 368

Db	129249	GCCTATTATTATTATTATTATTTTGGAGTGGAGTATTGCTTTGTCGCCAGGCCGA	129190
QY	369	GTGCAGTGGTGGATCTTGGCTCACTATAACCTCCACCTCCTGGTTCAAGTGATCCTCC	428
Db	129189	GTGCAGTGACGGGATCTCGGCCCACTGTAACTCTACCTCCTGGGTTCAAGTAATCTCC	129130
QY	429	CACTTTAGCCTCCTGAGTAGTGTGATTACAGGCGTGACCACCACTCCGGCTAAATTT	488
Db	129129	TACCTCAGCCTCCCGAGTAGCTGGGACTACAGGCACCCACCACTGGCTGATTTT	129070
QY	489	TGTATTTTATTAGAGACAGGGTTTCACCATGTGGCCAGGCTGTTCTCAAACCTCCTGA	548
Db	129069	TATATTTTAGTAGGGAC-GGGTTTCTCCATGTGGCCAGGCTGGTCTCAAACCTCCTGAC	129011
QY	549	CTCAAGGATCCGCCTGCCTCCACTTCCCAAGTCCCGAGATTACAGGTGTGAGTCACCA	608
Db	129010	CTCAAGTGATCTGCCTGCCTCAGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACCG	128951
QY	609	TGCCTGACCTTATA	622
Db	128950	CGACTGGCCAAAGA	128937

Search completed: September 30, 2004, 06:35:00
Job time : 14333 secs

This Page Blank (uspto)

Result No.	Score	Query ϵ		Length	DB	ID	Description
		Match					
1	3866	100.0		3866	9	US-09-967-768A-148	Sequence 148, App
2	3866	100.0		3866	13	US-10-342-887-899	Sequence 899, App
3	3866	100.0		3866	13	US-10-172-118-899	Sequence 899, App
4	3866	100.0		3866	15	US-10-081-119-13	Sequence 13, App
5	3866	100.0		3866	15	US-10-354-358-29	Sequence 29, App
6	3866	100.0		3866	15	US-10-116-712-663	Sequence 663, App
7	3866	100.0		3866	16	US-10-353-690-1	Sequence 1, Appl
8	3866	100.0		3866	16	US-10-366-288-15	Sequence 15, App
9	3866	100.0		3866	16	US-10-295-027-803	Sequence 803, App
10	3866	100.0		3866	16	US-10-295-027-832	Sequence 832, App
11	3866	100.0		3866	16	US-10-295-027-864	Sequence 864, App
12	2569.2	66.5		2574	16	US-10-295-027-481	Sequence 481, App
13	2569.2	66.5		2574	16	US-10-173-999-33	Sequence 33, App
14	402	10.4		2470	13	US-10-027-632-110237	Sequence 110237, App

QY 121 CTCAGCCTCTCAAGTAGCTGGGATTACAGGCATGTGCAACACCCCTGGCTAACTAATTT 180
Db 121 CTCAGCCTCTCAAGTAGCTGGGATTACAGGCATGTGCAACACCCCTGGCTAACTAATTT 180
QY 181 CTTTTCTATTAGTAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTTGAACCTCTG 240
Db 181 CTTTTCTATTAGTAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTTGAACCTCTG 240
QY 241 ACCTCAGGTGATCCACTTGCCTTGGCCTCCCAAGTGCTAGGATTACAGCCGTGAAAACGTG 300
Db 241 ACCTCAGGTGATCCACTTGCCTTGGCCTCCCAAGTGCTAGGATTACAGCCGTGAAAACGTG 300
QY 301 TGCCTGGCTGATCTTTTTTTTGTGTTGGATTTTGTGAACAGGGTCTCCCTTGGTCGCC 360
Db 301 TGCCTGGCTGATCTTTTTTTTGTGTTGGATTTTGTGAACAGGGTCTCCCTTGGTCGCC 360
QY 361 AGGCTGAGTGCAGTGGTGCGATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT 420
Db 361 AGGCTGAGTGCAGTGGTGCGATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT 420
QY 421 GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTAACAGCGTGCAACACCCCGG 480
Db 421 GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTAACAGCGTGCAACACCCCGG 480
QY 481 CTAATTTTGTATTTTATTAGAGACAGGGTTTACCATGTTGGCCAGGCTGTTCTCAAA 540
Db 481 CTAATTTTGTATTTTATTAGAGACAGGGTTTACCATGTTGGCCAGGCTGTTCTCAAA 540
QY 541 CTCCTGGACTCAAGGGATCCGCTGCCTCCACTTCCCAAGTCCCGAGATTACAGGTGTG 600
Db 541 CTCCTGGACTCAAGGGATCCGCTGCCTCCACTTCCCAAGTCCCGAGATTACAGGTGTG 600
QY 601 AGTCACCATGCCTGACCTTATAATTTCTTAAGTCATTTTCTGGTCCATTTCTTCTCTAG 660
Db 601 AGTCACCATGCCTGACCTTATAATTTCTTAAGTCATTTTCTGGTCCATTTCTTCTCTAG 660
QY 661 GGTCTCACAACAAATCTGCATTAGGCGGTACATAATCCTTAACCTCATGATTACACAAA 720
Db 661 GGTCTCACAACAAATCTGCATTAGGCGGTACATAATCCTTAACCTCATGATTACACAAA 720
QY 721 AGGAAGATGAAGTGATTTCATGATTTAGAAAGSGGAAGTAGTAAGCCCACTGCACACTCT 780
Db 721 AGGAAGATGAAGTGATTTCATGATTTAGAAAGSGGAAGTAGTAAGCCCACTGCACACTCT 780
QY 781 GGATGATGATCCTAAATCCAGATACAGTAATAATGGGGTATGGGAAGGTAGATAACAAA 840
Db 781 GGATGATGATCCTAAATCCAGATACAGTAATAATGGGGTATGGGAAGGTAGATAACAAA 840
QY 841 TTTTGGTTTAAATTAATTAATCTAATAATCTAATAATCTAATAATCTAATAATCTAATA 900
Db 841 TTTTGGTTTAAATTAATTAATCTAATAATCTAATAATCTAATAATCTAATAATCTAATA 900
QY 901 ATGTAAGACTGTACAGACTTCTTAGAAAACAGTTTGGGTCCATCTTTTCAATTTCCCCAG 960
Db 901 ATGTAAGACTGTACAGACTTCTTAGAAAACAGTTTGGGTCCATCTTTTCAATTTCCCCAG 960
QY 961 TGAGATTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAATTGATT 1020
Db 961 TGAGATTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAATTGATT 1020
QY 1021 CCATAATGAACAAAGTGAGAGACATTTAAATAAAGTTTAAATAAGAACCTTACTGATG 1080
Db 1021 CCATAATGAACAAAGTGAGAGACATTTAAATAAAGTTTAAATAAGAACCTTACTGATG 1080
QY 1081 AACTAAGCTTGAATAAATAATTTCTGCTGATACTACAGATAAATCGGGAACCTGTTAAACAAA 1140
Db 1081 AACTAAGCTTGAATAAATAATTTCTGCTGATACTACAGATAAATCGGGAACCTGTTAAACAAA 1140
QY 1141 TTATGATGATGGCAACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAA 1200
Db 1141 TTATGATGATGGCAACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAA 1200
QY 1201 ACAGTGTTCGCTAAGTGATGCTCTTTTAAATAAATTAAGTTGCTGTTACAGTCAAGCAA 1260

Db 1201 ACAGTGTTCGCTAAGTGATGCTCTTTTAAATAAATTAAGTTGCTGTTACAGTCAAGCAA 1260
QY 1261 TTGAAGCGCTTCCCCCAGATAAATAATGGCCAAATAGAGAGTTTGTGTAGAAATCAAGTGA 1320
Db 1261 TTGAAGCGCTTCCCCCAGATAAATAATGGCCAAATAGAGAGTTTGTGTAGAAATCAAGTGA 1320
QY 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG 1380
Db 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG 1380
QY 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACTGT 1440
Db 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACTGT 1440
QY 1441 CACAAGGTAATGTCAAAAAAGTAACAACCTTCTTCAAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAATGTCAAAAAAGTAACAACCTTCTTCAAAAAAGCTGTAGAACGTGGAGCAG 1500
QY 1501 TACCACCTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGC 1560
Db 1501 TACCACCTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAGCAGCTGC 1560
QY 1561 TTTTCAGAGGAGGAAAAAGAAATTTATCAGCATCTACGGTATTAACCTGCCCCAAGAAATCAT 1620
Db 1561 TTTTCAGAGGAGGAAAAAGAAATTTATCAGCATCTACGGTATTAACCTGCCCCAAGAAATCAT 1620
QY 1621 TTTCCGCTTCACTTGGGCAATTTACAGAAATAGGAAACAAACAGTTGTGATTCAGAGGACAGA 1680
Db 1621 TTTCCGCTTCACTTGGGCAATTTACAGAAATAGGAAACAAACAGTTGTGATTCAGAGGACAGA 1680
QY 1681 CTACTAAAGCCAGGTTTATATATGGAGAGAAATCCCAACAGATGCGCAATTTGGAAGAGTCC 1740
Db 1681 CTACTAAAGCCAGGTTTATATATGGAGAGAAATCCCAACAGATGCGCAATTTGGAAGAGTCC 1740
QY 1741 ACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCAATTTGGAAGAGTCC 1800
Db 1741 ACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCAATTTGGAAGAGTCC 1800
QY 1801 CAGTTAAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
QY 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAATGCCGAGATTTGGTGTGCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAATGCCGAGATTTGGTGTGCTGGATCTA 1920
QY 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGCTGTGATGAAAGAGTTCTGAACCTTATATTAATGATTCATATA 2040
Db 1981 TCAAGGAACCTCTGCTGTGATGAAAGAGTTCTGAACCTTATATTAATGATTCATATA 2040
QY 2041 CCTGAAGAATAAAACCGAATCAAGTCTTCTAGCTAAATTTAGAGAAACTAAAGAGTATC 2100
Db 2041 CCTGAAGAATAAAACCGAATCAAGTCTTCTAGCTAAATTTAGAGAAACTAAAGAGTATC 2100
QY 2101 AAGAACCCAGAGTTCCAGAGAGTAACCAAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACCCAGAGTTCCAGAGAGTAACCAAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
QY 2161 GTATTAAACCAAGATCTGCTGCTCATCTTCAAAATCACTGGCAGATTCGGGAGTTAGCCGAA 2220
Db 2161 GTATTAAACCAAGATCTGCTGCTCATCTTCAAAATCACTGGCAGATTCGGGAGTTAGCCGAA 2220
QY 2221 AAGTTAATAACAGAGCAAAACATACCACTTTTGGAGCAACCTGTCTTTTCAAGTTTCAAAAC 2280
Db 2221 AAGTTAATAACAGAGCAAAACATACCACTTTTGGAGCAACCTGTCTTTTCAAGTTTCAAAAC 2280
QY 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGAACCAAAATCTATTTTGTAAAGACACCAA 2340

Db 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTTGTAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACCTCCAGTTGTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACCTCCAGTTGTAAAGAAATGACT 2400
QY 2401 TTCCACCTGCTTGTGTCAGTTGTCAACACCTTATGGCCAACTGCCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGTCAGTTGTCAACACCTTATGGCCAACTGCCTGTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAAAATACCTTGCACCTCCACTTCAAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATACCTTGCACCTCCACTTCAAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
QY 2521 ATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATATAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATATAAATATG 2640
QY 2641 TGAACCTTAGAAGAACAGATACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAACAGATACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAAACACAGTGATAAGATCATCCGACTTTATGATTATGAATACACGG 2760
Db 2701 TGAATAAACTACAAACACAGTGATAAGATCATCCGACTTTATGATTATGAATACACGG 2760
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTTGATCTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTTGATCTTAATAGTTGGCTTAAAA 2820
QY 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
QY 2881 TTCACACAATCCATCAACATGGCATTGTTTCACAGTGATCTTAAACCAGCTAACCTTCTGA 2940
Db 2881 TTCACACAATCCATCAACATGGCATTGTTTCACAGTGATCTTAAACCAGCTAACCTTCTGA 2940
QY 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTGGGATTGCAACCAAATGCAACAGATA 3000
Db 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTGGGATTGCAACCAAATGCAACAGATA 3000
QY 3001 CAACAAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATATATGCCACCAGAACAA 3060
Db 3001 CAACAAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATATATGCCACCAGAACAA 3060
QY 3061 TCAAAGATATGTTCTCCTCCAGAGAGATGGGAAATCTAAGTCAAAAGATAAGCCCAAAA 3120
Db 3061 TCAAAGATATGTTCTCCTCCAGAGAGATGGGAAATCTAAGTCAAAAGATAAGCCCAAAA 3120
QY 3121 GTGATGTTTGGTCCCTTAGGATGTTATTTGTACTATATGACTTACGGGAAACACCATTTT 3180
Db 3121 GTGATGTTTGGTCCCTTAGGATGTTATTTGTACTATATGACTTACGGGAAACACCATTTT 3180
QY 3181 AGCAGATAATTAATCAGATTCTTAAATTACATGCCATAATTTGATCCTAATCATGAATTTG 3240
Db 3181 AGCAGATAATTAATCAGATTCTTAAATTACATGCCATAATTTGATCCTAATCATGAATTTG 3240
QY 3241 AATTTCCCGATATTTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCCGATATTTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
QY 3301 ACCCAAAACAGAGGATATCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db 3301 ACCCAAAACAGAGGATATCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
QY 3361 CTCATCCAGTTAACCAAAATGGCCCAAGGGACCACTGAAGAAATGAAAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAACCAAAATGGCCCAAGGGACCACTGAAGAAATGAAAATATGTTCTGGGCC 3420

QY 3421 AACTTGTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
Db 3421 AACTTGTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
QY 3481 ACTATAGTGGTGGTGAAGTCATAATTTCTTCATCCTCCAAGACTTTTGAAGAAAAAGGG 3540
Db 3481 ACTATAGTGGTGGTGAAGTCATAATTTCTTCATCCTCCAAGACTTTTGAAGAAAAAGGG 3540
QY 3541 GAAAAAATGATTTTGCAGTTATTTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600
Db 3541 GAAAAAATGATTTTGCAGTTATTTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600
QY 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAAACATCACTCTGAAGTGT 3660
Db 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAAACATCACTCTGAAGTGT 3660
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAAACTGTAAAAATAGCAACCAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAAACTGTAAAAATAGCAACCAC 3720
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTGTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
Db 3721 TTATGGCACTGTATATATTGTAGACTTGTGTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTCTAAAAAATTG 3840
Db 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTCTAAAAAATTG 3840
QY 3841 TAAATAAAGTTTTTGTGGCTAAAAATGA 3866
Db 3841 TAAATAAAGTTTTTGTGGCTAAAAATGA 3866

RESULT 2

US-10-342-887-899
; Sequence 899, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 50/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 50/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 899
; LENGTH: 3866
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-899

Query Match 100.0%; Score 3866; DB 13; Length 3866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATTCCTTTTTTTTTTTTTTTTGGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60
Db 1 GGAATTCCTTTTTTTTTTTTTTTTGGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60
QY 61 CAATGGCACAAATCTCAGCTTACTGCAACCTCCGCCTCCGGGTTCAAGCGATTCTCCTGTC 120

QY	2281	AGTCACCACCAATATCAACATCTAAATGGTTTGAACCCAAAATCTATTTGTAAAGACACCAA	2344
Db	2281		
QY	2281	AGTCACCACCAATATCAACATCTAAATGGTTTGAACCCAAAATCTATTTGTAAAGACACCAA	2344
Db	2281		
QY	2341	GCAGCAATACCTTGGATGATTACATGAGCTGTTTATAGAACTCCAGTTGTAAAGAATGACT	2400
Db	2341		
QY	2401	TTCCACCTGCTTGTCAGTTGTCAACACCTTATGGCCAACTGCCTGTTTCCAGCAGCAAC	2460
Db	2401		
QY	2461	AGCATCAAAATACTTGCACCTCCACTTCAAAAATTTACAGGTTTTAGCATCTTCTTCCAGCAA	2520
Db	2461		
QY	2521	ATGAATGCATTTTCGGTTAAAGGAAGAATTTATTCCATATTAAGCAGATAGGAAGTGGAG	2580
Db	2521		
QY	2581	GTTCAAGCAAGGTATTTACAGGTGTTAAATGAAAAGAAAACAGATATATGCTATAAAATATG	2640
Db	2581		
QY	2641	TGAACCTTAGAAGACAGATAACCAAACCTCTTGATAGTTACCGGAACGAAATAGCTTATT	2700
Db	2641		
QY	2701	TGAATAAACTCAACAACACAGTGATAAGATCATCCGACTTTATGATTATGAAATCACGG	2760
Db	2701		
QY	2761	ACCAGTACATCTACATGGTAAATGGAGTGTGGAATAATTGATCTTAATAGTTGGCTTAAAA	2820
Db	2761		
QY	2821	AGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG	2880
Db	2821		
QY	2881	TTCACACAATCCATCAACATGGCATTGTTCCAGTGATCTTAAACCAGCTAACITTCTGA	2940
Db	2881		
QY	2941	TAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTGCAAAACCAAATGCCAACCCAGATA	3000
Db	2941		
QY	3001	CAACAAGTGTGTTAAAGATTCTCAGGTTGGCACAGTTAATTATATGCCACCAGAAAGCAA	3060
Db	3001		
QY	3061	TCAAAGATATGCTTCCCTCCAGAGAGAATGGGAAATCTAAGTCAAAGATAAGCCCCAAAA	3120
Db	3061		
QY	3121	GTGATGTTTGGTCCCTTAGGATGTAATTTGTACTATATGACTTACGGGAAAAACACCATTTTC	3180
Db	3121		
QY	3181	AGCAGATAATTAAATCAGATTTCTAAATTACATGCCATAATTGATCCTAATCATGAATTG	3240
Db	3181		
QY	3241	AATTTCCCGATATTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG	3300
Db	3241		
QY	3301	ACCCAAAAACAGAGGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA	3360
Db	3301		

QY	3361	CTCATCCAGTTAAACCAAAATGGCCAAAGGGAACCACTGAAGAAAATGAAAATATGTTCTGGGCC	3420
Db	3361	CTCATCCAGTTAAACCAAAATGGCCAAAGGGAACCACTGAAGAAAATGAAAATATGTTCTGGGCC	3420
QY	3421	AACCTGTTGGTCTGAATTTCTCCTAACTCCATTTTGAAAGCTGCTAAAACTTTATATGAAC	3480
Db	3421	AACCTGTTGGTCTGAATTTCTCCTAACTCCATTTTGAAAGCTGCTAAAACTTTATATGAAC	3480
QY	3481	ACTATAGTGGTGGGAAAGTCATAAATCTTCATCCTCCAAGACTTTTGAAAAAAAAGGG	3540
Db	3481	ACTATAGTGGTGGGAAAGTCATAAATCTTCATCCTCCAAGACTTTTGAAAAAAAAGGG	3540
QY	3541	GAAAAAAATGATTTGCAGTTATTTCGTAATGTCAGATAGGAGGTATAAAAATATATTGGACT	3600
Db	3541	GAAAAAAATGATTTGCAGTTATTTCGTAATGTCAGATAGGAGGTATAAAAATATATTGGACT	3600
QY	3601	GTTATACCTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAAGTGTT	3660
Db	3601	GTTATACCTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAAGTGTT	3660
QY	3661	ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAACGTAAAAATAGCAACCAC	3720
Db	3661	ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAACGTAAAAATAGCAACCAC	3720
QY	3721	TTATGGCACTGTATATATTGTAGACTTTGTTTCTCTGTTTTATGCTCTTGTGTAATCTAC	3780
Db	3721	TTATGGCACTGTATATATTGTAGACTTTGTTTCTCTGTTTTATGCTCTTGTGTAATCTAC	3780
QY	3781	TTGACATCATTTTACTCTTTGGAATAGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTG	3840
Db	3781	TTGACATCATTTTACTCTTTGGAATAGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTG	3840
QY	3841	TAAATAAAGTTTTGTGGCTAAAATGA	3866
Db	3841	TAAATAAAGTTTTGTGGCTAAAATGA	3866

RESULT 3

```

US-10-172-118-899
; Sequence 899, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 899
; LENGTH: 3866
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003318
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-899

```

Query Match	100.0%;	Score 3866;	DB 13;	Length 3866;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3866;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy

1 GGAAATCCCTTTT TTTTTTTTTTTT TTGAGATGGAGTTTCAC TCTGTGTGGCCAGGCTGGAGTG 60

Dβ

1 GGAAATCCCTTTT TTTTTTTTTTTT TTGAGATGGAGTTTCAC TCTGTGTGGCCAGGCTGGAGTG 60

QY 61 CAATGGCACAACTCTCAGCTTACTGCAACCTCCGCTCCGGGTTTCAAGCGATTCTCCTGC 120
Db |||||
61 CAATGGCACAACTCTCAGCTTACTGCAACCTCCGCTCCGGGTTTCAAGCGATTCTCCTGC 120
QY 121 CTCAGCCTCTCAAGTAGCTGGGATTACAGGCATGTGCACACCCCTGGCTAACTAATTT 180
Db |||||
121 CTCAGCCTCTCAAGTAGCTGGGATTACAGGCATGTGCACACCCCTGGCTAACTAATTT 180
QY 181 CTTTCTATTAGTAGAGATGGGTTTCAACCATGTTGGTCAGGCTGGTCTTGAACCTCCTG 240
Db |||||
181 CTTTCTATTAGTAGAGATGGGTTTCAACCATGTTGGTCAGGCTGGTCTTGAACCTCCTG 240
QY 241 ACCTCAGGTGATCCACTTGCCTTGGCCTCCCAAGTGTCTAGGATTACAGCGGTGAACCTG 300
Db |||||
241 ACCTCAGGTGATCCACTTGCCTTGGCCTCCCAAGTGTCTAGGATTACAGCGGTGAACCTG 300
QY 301 TGCCTGGCTGATCTCTTTTGTGTTGGATTTTGAACAGGGTCTCCCTTGGTCGCC 360
Db |||||
301 TGCCTGGCTGATCTCTTTTGTGTTGGATTTTGAACAGGGTCTCCCTTGGTCGCC 360
QY 361 AGGCTGGAGTGCAGTGGTGGATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT 420
Db |||||
361 AGGCTGGAGTGCAGTGGTGGATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT 420
QY 421 GATCCTCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGCACACACACCCGG 480
Db |||||
421 GATCCTCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGCACACACACCCGG 480
QY 481 CTAATTTTCTATTTTATTAGAGACAGGGTTTCAACATGTTGGCCAGGCTGTTCTCAAA 540
Db |||||
481 CTAATTTTCTATTTTATTAGAGACAGGGTTTCAACATGTTGGCCAGGCTGTTCTCAAA 540
QY 541 CTCCTGGACTCAAGGGATCCGCTGCCCTCCAATTCCCAAGTCCCGAGATTACAGGTGTG 600
Db |||||
541 CTCCTGGACTCAAGGGATCCGCTGCCCTCCAATTCCCAAGTCCCGAGATTACAGGTGTG 600
QY 601 AGTCACCATGCTGACCTTAATAATCTTAAGTCAATTTTCTGGTCCATTTCTCCTTAG 660
Db |||||
601 AGTCACCATGCTGACCTTAATAATCTTAAGTCAATTTTCTGGTCCATTTCTCCTTAG 660
QY 661 GGTCTCACAACAAATCTGCATTAGCGGTACAAATAATCCTTAACCTTCAATGATTACAAA 720
Db |||||
661 GGTCTCACAACAAATCTGCATTAGCGGTACAAATAATCCTTAACCTTCAATGATTACAAA 720
QY 721 AGGAAGATGAAGTGAATCATGATTTAGAAAGGGGAGTAGTAAGCCCACTGCACACTCCT 780
Db |||||
721 AGGAAGATGAAGTGAATCATGATTTAGAAAGGGGAGTAGTAAGCCCACTGCACACTCCT 780
QY 781 GGATGATGATCCTAAATCCAGATACAGTAAATGGGATGGGATGGGAAGTAGAATAACAAA 840
Db |||||
781 GGATGATGATCCTAAATCCAGATACAGTAAATGGGATGGGATGGGAAGTAGAATAACAAA 840
QY 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAACATTTTGGATACATGTTGATGTA 900
Db |||||
841 TTTGGTTTAAATTAATTAATCTAAATATCTAAACATTTTGGATACATGTTGATGTA 900
QY 901 ATGTAAGACTGTACAGACTTCTAGAAAACAGTTGGGTTCCATCTTTTCATTTCCCCAG 960
Db |||||
901 ATGTAAGACTGTACAGACTTCTAGAAAACAGTTGGGTTCCATCTTTTCATTTCCCCAG 960
QY 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGATTGACAATTGATT 1020
Db |||||
961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGATTGACAATTGATT 1020
QY 1021 CCATAATGAACAAAGTGAGAGACATTAATAAATAGTTTAAATAAGAGACCTTACTGATG 1080
Db |||||
1021 CCATAATGAACAAAGTGAGAGACATTAATAAATAGTTTAAATAAGAGACCTTACTGATG 1080
QY 1081 AACTAAGCTTGAATAAAATTTCTGCTGATCTACAGATAAATCGGGAACTGTTAAACCAA 1140
Db |||||
1081 AACTAAGCTTGAATAAAATTTCTGCTGATCTACAGATAAATCGGGAACTGTTAAACCAA 1140

QY 1141 TTATGATGATGGCAAAACAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAA 1200
Db |||||
1141 TTATGATGATGGCAAAACAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAA 1200
QY 1201 ACAGTGTTCGCTAAGTAGTGTCTTTTAAATAAATGATTTGGTCTGTTACAGTCAAGCAA 1260
Db |||||
1201 ACAGTGTTCGCTAAGTAGTGTCTTTTAAATAAATGATTTGGTCTGTTACAGTCAAGCAA 1260
QY 1261 TTGAAGCGCTTCCCCCAGATAAATATGGCCAAATGAGAGTTTGTGCTAGAAATCAAGTGA 1320
Db |||||
1261 TTGAAGCGCTTCCCCCAGATAAATATGGCCAAATGAGAGTTTGTGCTAGAAATCAAGTGA 1320
QY 1321 GATTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG 1380
Db |||||
1321 GATTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG 1380
QY 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACCTGT 1440
Db |||||
1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACCTGT 1440
QY 1441 CACAAGGTAATGTCAAAAAAAGTAAACAACCTTCTCAAAAAAGCTGTAGAACGTGGAGCAG 1500
Db |||||
1441 CACAAGGTAATGTCAAAAAAAGTAAACAACCTTCTCAAAAAAGCTGTAGAACGTGGAGCAG 1500
QY 1501 TACCACCTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
Db |||||
1501 TACCACCTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
QY 1561 TTTTCAGAGGAGGAAAGAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
Db |||||
1561 TTTTCAGAGGAGGAAAGAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
QY 1621 TTTCCGGTTCACTTGGGCAATTTACAGAAATAGGAACAACAGTTGTGATTTCCAGAGGACAGA 1680
Db |||||
1621 TTTCCGGTTCACTTGGGCAATTTACAGAAATAGGAACAACAGTTGTGATTTCCAGAGGACAGA 1680
QY 1681 CTACTAAAGCCAGGTTTTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTT 1740
Db |||||
1681 CTACTAAAGCCAGGTTTTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTT 1740
QY 1741 ACCGGAATTCATTGAGACAAACTAACAAACCTAACAACAGTCATGCCCATTTGGAAGAGTCC 1800
Db |||||
1741 ACCGGAATTCATTGAGACAAACTAACAAACCTAACAACAGTCATGCCCATTTGGAAGAGTCC 1800
QY 1801 CAGTTAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db |||||
1801 CAGTTAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
QY 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCTGGATCTA 1920
Db |||||
1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCTGGATCTA 1920
QY 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
Db |||||
1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGGTGTGATGAAAAGAGTCTGAACTTATTAATTAATGATTTCAATAA 2040
Db |||||
1981 TCAAGGAACCTCTGGTGTGATGAAAAGAGTCTGAACTTATTAATTAATGATTTCAATAA 2040
QY 2041 CCCTGAAGAAATAAAACCGGAATCAAGTCTTCTAGTAAATTAGAAGAACTAAAGAGTATC 2100
Db |||||
2041 CCCTGAAGAAATAAAACCGGAATCAAGTCTTCTAGTAAATTAGAAGAACTAAAGAGTATC 2100
QY 2101 AAGAACACAGAGGTTCCAGAGAGTAACCCAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db |||||
2101 AAGAACACAGAGGTTCCAGAGAGTAACCCAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
QY 2161 GTATTAAACCAAGATCCTGCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Db |||||
2161 GTATTAAACCAAGATCCTGCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGGCAACCTGTCTTTTTCAGTTTCAAAAC 2280

Db	1	GGAATTCCTTTT	TTTTTTTGAGATGGAGTTTCACACTCTGTGTTGGCCAGGCTGGAGTG	60
QY	61	CAATGGCACAAATCTCAGCTTA	CTGCCAACCTCGGCCTCCGGGTTTCAAGCGATTCTCCTGC	120
Db	61	CAATGGCACAAATCTCAGCTTA	CTGCCAACCTCGGCCTCCGGGTTTCAAGCGATTCTCCTGC	120
QY	121	CTCAGCCTCTCAAGTAGCTGGG	AATACAGGCATGTGCCACCACCCCCTGGCTAAGTAATAATTT	180
Db	121	CTCAGCCTCTCAAGTAGCTGGG	AATACAGGCATGTGCCACCACCCCCTGGCTAAGTAATAATTT	180
QY	181	CTTTTCTAATTTAGTAGATGGG	TTCACCATGTTGGTCAGGCTGGTCTTGAACCTCCTG	240
Db	181	CTTTTCTAATTTAGTAGATGGG	TTCACCATGTTGGTCAGGCTGGTCTTGAACCTCCTG	240
QY	241	ACCTCAGTGCATCCACTTGCCT	TGGCCTCCCAAAGTGCTAGGATTACAGCCGTGAAA	300
Db	241	ACCTCAGTGCATCCACTTGCCT	TGGCCTCCCAAAGTGCTAGGATTACAGCCGTGAAA	300
QY	301	TGCCTGGCTGATTCTTTT	TGTTGTTGGATTTTGTAAACAGGGTCTCCCTGGTTCGCC	360
Db	301	TGCCTGGCTGATTCTTTT	TGTTGTTGGATTTTGTAAACAGGGTCTCCCTGGTTCGCC	360
QY	361	AGGCTGGAGTGCAGTGGTGCG	ATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT	420
Db	361	AGGCTGGAGTGCAGTGGTGCG	ATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT	420
QY	421	GATCCTCCCACTTTAGCCTC	CTGAGTAGCTGTGATTACAGGCGTGCCACACACACCCGG	480
Db	421	GATCCTCCCACTTTAGCCTC	CTGAGTAGCTGTGATTACAGGCGTGCCACACACACCCGG	480
QY	481	CTAATTTTGTATTTT	TATTAGACACAGGGTTTCACCATGTTGGCCAGGCTGTTCTCAA	540
Db	481	CTAATTTTGTATTTT	TATTAGACACAGGGTTTCACCATGTTGGCCAGGCTGTTCTCAA	540
QY	541	CTCCTGGACTCAAGGGATCCG	CGCTGCCCTCCACTTCCCCAAAGTCCCGAGATTACAGGTGTG	600
Db	541	CTCCTGGACTCAAGGGATCCG	CGCTGCCCTCCACTTCCCCAAAGTCCCGAGATTACAGGTGTG	600
QY	601	AGTCACCATGCCTGACCTT	ATAATTTCTTAAGTCATTTTTTCTGGTCCATTTCTTCCTTAG	660
Db	601	AGTCACCATGCCTGACCTT	ATAATTTCTTAAGTCATTTTTTCTGGTCCATTTCTTCCTTAG	660
QY	661	GGTCTCACAACAAATCTGCA	TATAGCGGTACAATAATCTCTTAACCTCATGATTACAAAA	720
Db	661	GGTCTCACAACAAATCTGCA	TATAGCGGTACAATAATCTCTTAACCTCATGATTACAAAA	720
QY	721	AGGAAGATGAAGTGATT	CATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT	780
Db	721	AGGAAGATGAAGTGATT	CATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT	780
QY	781	GGATGATGATCCTAAATCC	AGATACAGTAAAAATGGGGTATGGGAAGGTAGAATAACAAAA	840
Db	781	GGATGATGATCCTAAATCC	AGATACAGTAAAAATGGGGTATGGGAAGGTAGAATAACAAAA	840
QY	841	TTTGGTTTAAATTAAT	TATCTFAAATATCTAAAAACATTTTTTGGATACATTTGTGATGTA	900
Db	841	TTTGGTTTAAATTAAT	TATCTFAAATATCTAAAAACATTTTTTGGATACATTTGTGATGTA	900
QY	901	ATGTAAGACTGTACAGACT	CTCCTAGAAAACAGTTTGGGTTCCCATCTTTTTCATTTCCCCCAG	960
Db	901	ATGTAAGACTGTACAGACT	CTCCTAGAAAACAGTTTGGGTTCCCATCTTTTTCATTTCCCCCAG	960
QY	961	TGCAGTTTCTGTAGAAATG	GAATCCGAGGATTTAAGTGGCAGAGAAATTGACAATTGATT	1020
Db	961	TGCAGTTTCTGTAGAAATG	GAATCCGAGGATTTAAGTGGCAGAGAAATTGACAATTGATT	1020
QY	1021	CCATAATGAACAAAGTG	AGACATTTAAAAATTAAGTTTAAAAATGAAGACCTTACTGATG	1080
Db	1021	CCATAATGAACAAAGTG	AGACATTTAAAAATTAAGTTTAAAAATGAAGACCTTACTGATG	1080
QY	1081	AACTAAGCTTGAATAAAA	TTTTCTGCTGATACTACAGATAAAGTTCGGGAAGTGTTAACCAAA	1140

Db	1081	AAC	TAA	GCTT	GAA	TAAAAA	TTCT	GCTG	ATACT	ACAG	ATAA	CTCG	GAACT	GTTAA	ACCA	AAA	114															
QY	1141	TTA	TGAT	GATG	GCAAA	CAAC	CC	CAG	AGG	ACTG	GTG	TTG	TTG	TGCT	CAAA	CTAG	AGAAAA	1200														
Db	1141	TTA	TGAT	GATG	GCAAA	CAAC	CC	CAG	AGG	ACTG	GTG	TTG	TTG	TGCT	CAAA	CTAG	AGAAAA	1200														
QY	1201	ACA	GTG	TTCC	GCTA	AGT	GAT	GCT	CTTT	TAA	TAA	ATTT	GAT	TGG	T	CGTT	ACA	GCTCAAGCAA	1260													
Db	1201	ACA	GTG	TTCC	GCTA	AGT	GAT	GCT	CTTT	TAA	TAA	ATTT	GAT	TGG	T	CGTT	ACA	GCTCAAGCAA	1260													
QY	1261	TTG	AAG	CGCT	TTCC	CC	CAG	ATAA	TAT	TG	CC	CAAA	TG	AGAG	TTT	TGCT	AG	AAATTC	CAAGTGA	1320												
Db	1261	TTG	AAG	CGCT	TTCC	CC	CAG	ATAA	TAT	TG	CC	CAAA	TG	AGAG	TTT	TGCT	AG	AAATTC	CAAGTGA	1320												
QY	1321	GAT	TTG	CTGA	TTAA	AA	GCT	ATT	CA	AG	AG	CC	CAG	TG	AT	GC	AC	GT	TACT	TTCAAATGG	1380											
Db	1321	GAT	TTG	CTGA	TTAA	AA	GCT	ATT	CA	AG	AG	CC	CAG	TG	AT	GC	AC	GT	TACT	TTCAAATGG	1380											
QY	1381	CC	AG	AGCA	AACT	GC	AA	AGAA	ATTT	TG	CTTT	TG	TT	CA	TAT	AT	CTTT	TG	CA	CAATTT	GAACTGT	1440										
Db	1381	CC	AG	AGCA	AACT	GC	AA	AGAA	ATTT	TG	CTTT	TG	TT	CA	TAT	AT	CTTT	TG	CA	CAATTT	GAACTGT	1440										
QY	1441	CAC	AA	GGT	AT	TG	CA	AAAA	AA	GT	AA	CA	AACT	CT	TT	CA	AAAA	AGCT	GT	AG	AA	CGTGG	AGCAG	1500								
Db	1441	CAC	AA	GGT	AT	TG	CA	AAAA	AA	GT	AA	CA	AACT	CT	TT	CA	AAAA	AGCT	GT	AG	AA	CGTGG	AGCAG	1500								
QY	1501	TAC	CA	CT	AG	AA	AT	GC	TG	GA	AA	TTG	CC	CT	GC	CG	AA	TTT	TAA	AC	CCT	CC	AAAA	AAAGCAG	CTGC	1560						
Db	1501	TAC	CA	CT	AG	AA	AT	GC	TG	GA	AA	TTG	CC	CT	GC	CG	AA	TTT	TAA	AC	CCT	CC	AAAA	AAAGCAG	CTGC	1560						
QY	1561	TTT	C	AG	AG	G	AG	AAAA	AA	GA	AA	TT	T	AT	C	AG	CA	T	T	A	C	T	A	CT	GC	CC	CAAG	AAATCAT	1620			
Db	1561	TTT	C	AG	AG	G	AG	AAAA	AA	GA	AA	TT	T	AT	C	AG	CA	T	T	A	C	T	A	CT	GC	CC	CAAG	AAATCAT	1620			
QY	1621	TTT	CC	CG	TT	CA	CT	TG	GC	CA	TTT	T	AC	AG	AA	TAG	GA	AA	CA	AC	AG	TTG	T	G	ATT	CC	AG	AG	GAC	AGA	1680	
Db	1621	TTT	CC	CG	TT	CA	CT	TG	GC	CA	TTT	T	AC	AG	AA	TAG	GA	AA	CA	AC	AG	TTG	T	G	ATT	CC	AG	AG	GAC	AGA	1680	
QY	1681	CT	ACT	AA	AG	CC	AG	GT	TTT	T	AT	TG	AG	AG	AA	CA	CT	GC	CC	CA	AG	CT	AT	GC	CC	CA	AG	AA	AT	AG	GT	1740
Db	1681	CT	ACT	AA	AG	CC	AG	GT	TTT	T	AT	TG	AG	AG	AA	CA	CT	GC	CC	CA	AG	CT	AT	GC	CC	CA	AG	AA	AT	AG	GT	1740
QY	1741	ACC	GG	AA	TT	CA	TT	G	AG	AC	AA	CT	AA	CA	AG	CT	AT	GC	CC	CA	AG	CT	AT	GC	CC	CA	AG	AG	AG	AG	TC	1800
Db	1741	ACC	GG	AA	TT	CA	TT	G	AG	AC	AA	CT	AA	CA	AG	CT	AT	GC	CC	CA	AG	CT	AT	GC	CC	CA	AG	AG	AG	AG	TC	1800
QY	1801	CAG	TTA	AC	CT	T	CT	AA	TAG	CC	CAG	ATT	G	TG	AT	GT	GA	AG	AC	AG	AT	G	AT	CA	G	T	T	G</				

QY 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
QY 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTAAGACACCAA 2340
Db 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAGAAATGACT 2400
QY 2401 TTCCACCTGTTGTCAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGTTGTCAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAAATACCTGCCACTCCACTTCAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAATACCTGCCACTCCACTTCAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
QY 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGTATTTCAAGGTGTTAAATGAAAAGAAAACAGATATATGCTATAAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTCAAGGTGTTAAATGAAAAGAAAACAGATATATGCTATAAAATATG 2640
QY 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTTATGATTATGAAATCACGG 2760
Db 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTTATGATTATGAAATCACGG 2760
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCITTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCITTAATAGTTGGCTTAAAA 2820
QY 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
QY 2881 TTCACACAATCCATCAACATGGCATTGTTTCAAGTGATCTTAAACCAGCTAACCTTCTGA 2940
Db 2881 TTCACACAATCCATCAACATGGCATTGTTTCAAGTGATCTTAAACCAGCTAACCTTCTGA 2940
QY 2941 TAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTGCAACCAAAATGCAACCAGATA 3000
Db 2941 TAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTGCAACCAAAATGCAACCAGATA 3000
QY 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGATTAAATATATGCCACCAAGAGCAA 3060
Db 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGATTAAATATATGCCACCAAGAGCAA 3060
QY 3061 TCAAAGATATGCTTCTCCAGAGAGAATGGGAAATCTAAGTCAAAGATAAGCCCCCAAAA 3120
Db 3061 TCAAAGATATGCTTCTCCAGAGAGAATGGGAAATCTAAGTCAAAGATAAGCCCCCAAAA 3120
QY 3121 GTGATGTTTGGTCCCTTAGGATGTTATTTGTACTATPATGACTTACGGGAAAAACACCATTTT 3180
Db 3121 GTGATGTTTGGTCCCTTAGGATGTTATTTGTACTATPATGACTTACGGGAAAAACACCATTTT 3180
QY 3181 AGCAGATAATTAATCAGATTTCTAAATTACATGCCATAATTGATCCCTAATCATGAAATTG 3240
Db 3181 AGCAGATAATTAATCAGATTTCTAAATTACATGCCATAATTGATCCCTAATCATGAAATTG 3240
QY 3241 AATTTCCGATATTCAGAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCGATATTCAGAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300

QY 3301 ACCCAAAACAGAGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db 3301 ACCCAAAACAGAGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
QY 3361 CTCATCCAGTTAACCAAAATGGCCAAGGGAACCACTGAAGAAATGAAAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAACCAAAATGGCCAAGGGAACCACTGAAGAAATGAAAATATGTTCTGGGCC 3420
QY 3421 AACTTGTGTTGGTCTGAATTCCTCTAACTCCATTTTGAAGCTGCTTAAACTTTTATATGAAC 3480
Db 3421 AACTTGTGTTGGTCTGAATTCCTCTAACTCCATTTTGAAGCTGCTTAAACTTTTATATGAAC 3480
QY 3481 ACTATAGTGGTGGTGAAGTCTAATACTCTCAAGACTTTTGAAGCTTTTGAAGCTTTTGAAGGG 3540
Db 3481 ACTATAGTGGTGGTGAAGTCTAATACTCTCAAGACTTTTGAAGCTTTTGAAGCTTTTGAAGGG 3540
QY 3541 GAAAAAAATGATTGCAATTTCGTAATTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600
Db 3541 GAAAAAAATGATTGCAATTTCGTAATTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600
QY 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAACTGTAAAAATAGCAACCAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAACTGTAAAAATAGCAACCAC 3720
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTCTTCTCTGTTTATGCTCTGTTGTAATCTAC 3780
Db 3721 TTATGGCACTGTATATATTGTAGACTTGTCTTCTCTGTTTATGCTCTGTTGTAATCTAC 3780
QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGTAGCAAGTATATTCTAAAAAACTTTG 3840
Db 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGTAGCAAGTATATTCTAAAAAACTTTG 3840
QY 3841 TAAATAAAGTTTTGTGGCTAAAAATGA 3866
Db 3841 TAAATAAAGTTTTGTGGCTAAAAATGA 3866

RESULT 5

US-10-354-358-29
; Sequence 29, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2150,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 32230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MPI02-020P1RNMNIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10

Db 1741 ACCGGAATTCATTGAGACAAAACCTAACAAAACCTAACAGTCATGCCCATTTGGAAGAGTCC 1800
QY 1801 CAGTTAACCTTCTAAATAGCCCAAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1850
Db 1801 CAGTTAACCTTCTAAATAGCCCAAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1850
QY 1861 GTTTTATGAAGAACAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCCTGGATCTA 1920
Db 1861 GTTTTATGAAGAACAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCCTGGATCTA 1920
QY 1921 AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGGTGTGATGAAAGAGTTCCTGAACCTTATTTACTGATTCAAATAA 2040
Db 1981 TCAAGGAACCTCTGGTGTGATGAAAGAGTTCCTGAACCTTATTTACTGATTCAAATAA 2040
QY 2041 CCTGGAAGAAATAAAGCGGAATCAAGTCTCTAGCTAAATTAAGAAACCTAAAGAGTATC 2100
Db 2041 CCTGGAAGAAATAAAGCGGAATCAAGTCTCTAGCTAAATTAAGAAACCTAAAGAGTATC 2100
QY 2101 AAGAACCCAGAGGTTCCAGAGAGTAACCCAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACCCAGAGGTTCCAGAGAGTAACCCAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
QY 2161 GTATTAAACCAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACCAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAAGAAACATACCACCTTTTGAGCAACCTGCTTTTCAGTTTCAAAAAC 2280
Db 2221 AAGTTAATACAGAGCAAGAAACATACCACCTTTTGAGCAACCTGCTTTTCAGTTTCAAAAAC 2280
QY 2281 AGTCACCACCAATATCAACATCTAAATGCTTTGACCCAAAATCTATTGTAAGACACCAA 2340
Db 2281 AGTCACCACCAATATCAACATCTAAATGCTTTGACCCAAAATCTATTGTAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACT 2400
QY 2401 TTCCACCTGCTTGTGATGTCAACACCTTATGGCCAAACCTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGATGTCAACACCTTATGGCCAAACCTGCTGTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAAATACCTTGCCACTCCACTTCAAAATTTACAGGTTTTCAGCATCTTCTTACGCAA 2520
Db 2461 AGCATCAAATACCTTGCCACTCCACTTCAAAATTTACAGGTTTTCAGCATCTTCTTTCAGCAA 2520
QY 2521 ATGAATGCATTTTCGGTTAAAGGAAGAAATTTATCCATATTAAAGCAGATAGGAAGTGAG 2580
Db 2521 ATGAATGCATTTTCGGTTAAAGGAAGAAATTTATCCATATTAAAGCAGATAGGAAGTGAG 2580
QY 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAGAAAGAAACAGATATATGCTATAAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAGAAAGAAACAGATATATGCTATAAAATATG 2640
QY 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCAGG 2760
Db 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCAGG 2760
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTTAATAGTTGGCTTAAAA 2820
QY 2821 AGAAAAATCCATTGATCCATGGGAACCGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAATCCATTGATCCATGGGAACCGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880

QY 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Db 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAGCTAACTTTCTGA 2940
QY 2941 TAGTTGATGGAAATGCTAAAGCTAAATGATTTTGGGATTTGCAAAACCAAAATGCAACCCAGATA 3000
Db 2941 TAGTTGATGGAAATGCTAAAGCTAAATGATTTTGGGATTTGCAAAACCAAAATGCAACCCAGATA 3000
QY 3001 CAACAAAGTGTGTTAAAGATTTCTCAGGTGGCACAGTTAATATATATGCCACCAGAAAGCAA 3060
Db 3001 CAACAAAGTGTGTTAAAGATTTCTCAGGTGGCACAGTTAATATATATGCCACCAGAAAGCAA 3060
QY 3061 TCAAAAGATATGTTCTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
Db 3061 TCAAAAGATATGTTCTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
QY 3121 GTGATGTTTGGTCCCTTAGGATGTTTGTGATCTATATGACTTACGGGAAAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCCCTTAGGATGTTTGTGATCTATATGACTTACGGGAAAAACACCATTTTC 3180
QY 3181 AGCAGATAATTAATCAGATTTCTAAATTAACATGCCATAATTAATGATCCTAATCATGAAATTC 3240
Db 3181 AGCAGATAATTAATCAGATTTCTAAATTAACATGCCATAATTAATGATCCTAATCATGAAATTC 3240
QY 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGG 3300
Db 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGG 3300
QY 3301 ACCAAAAACAGAGGATATCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db 3301 ACCAAAAACAGAGGATATCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
QY 3361 CTGATCCAGTTAACCAATGGCCAAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
Db 3361 CTGATCCAGTTAACCAATGGCCAAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
QY 3421 AACTTGTGTTGGTCTGAATTTCTTAACCTCCATTTTGAAGCTGCTTAAACTTTTATATGAAC 3480
Db 3421 AACTTGTGTTGGTCTGAATTTCTTAACCTCCATTTTGAAGCTGCTTAAACTTTTATATGAAC 3480
QY 3481 ACTATAGTGGTGGTGAAGTCAATAATTTCTTCATCTCCTCAAGACTTTTGAAAAAAAAGGG 3540
Db 3481 ACTATAGTGGTGGTGAAGTCAATAATTTCTTCATCTCCTCAAGACTTTTGAAAAAAAAGGG 3540
QY 3541 GAAAAAAATGATTTGCGATTTATTCGTAATGTGATAGGAGGTATAAAAATATATTTGGACT 3600
Db 3541 GAAAAAAATGATTTGCGATTTATTCGTAATGTGATAGGAGGTATAAAAATATATTTGGACT 3600
QY 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAAACATCACTCTGAAGTGT 3660
Db 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAAACATCACTCTGAAGTGT 3660
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTTATCTTTAAAAAGAAAACTGTAAAAATAGCAACCCAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTTATCTTTAAAAAGAAAACTGTAAAAATAGCAACCCAC 3720
QY 3721 TTATGSCACTGTATATATTTGAGACTTGTGTTTCTGTTTATGCTCTGTTGTAATCTAC 3780
Db 3721 TTATGSCACTGTATATATTTGAGACTTGTGTTTCTGTTTATGCTCTGTTGTAATCTAC 3780
QY 3781 TTGACATCATTTTACTCTTGAATAGTGGTGGATAGCAAGTATATTTCTAAAAAACTTTT 3840
Db 3781 TTGACATCATTTTACTCTTGAATAGTGGTGGATAGCAAGTATATTTCTAAAAAACTTTT 3840
QY 3841 TAAATAAAGTTTTTGTGGCTAAAAATGA 3866
Db 3841 TAAATAAAGTTTTTGTGGCTAAAAATGA 3866

```

; Publication No. US20030194764A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Switzer, Ann
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.568
; CURRENT APPLICATION NUMBER: US/10/116,712
; CURRENT FILING DATE: 2002-04-07
; NUMBER OF SEQ ID NOS: 670
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 663
; LENGTH: 3866
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-712-663

Query Match      100.0%; Score 3866; DB 15; Length 3866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGAATTCCTTTTCTTTTTCAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60
DB      1  GGAATTCCTTTTCTTTTTCAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60

QY      61  CAATGGCACAATCTCAGCTTACTGCAACCTCCGCCCTCCGGGTTCAAGCGATTCTCCTGC 120
DB      61  CAATGGCACAATCTCAGCTTACTGCAACCTCCGCCCTCCGGGTTCAAGCGATTCTCCTGC 120

QY      121  CTCAGCCCTCAAGTAGCTGGGATTACAGGCATGTGCCACCACCCCTGGCTAACTAATTT 180
DB      121  CTCAGCCCTCAAGTAGCTGGGATTACAGGCATGTGCCACCACCCCTGGCTAACTAATTT 180

QY      181  CTTTCTAATTAGTAGAGATGGGGTTTCAACCATGTGGTCAGGCTGGTCTTGAACCTCGT 240
DB      181  CTTTCTAATTAGTAGAGATGGGGTTTCAACCATGTGGTCAGGCTGGTCTTGAACCTCGT 240

QY      241  ACCTCAGTGATCCACTTGCCTTGGCCTCCCAAAGTGCTAGGATTACAGCCGTGAAACTG 300
DB      241  ACCTCAGTGATCCACTTGCCTTGGCCTCCCAAAGTGCTAGGATTACAGCCGTGAAACTG 300

QY      301  TGCCTGGCTGATCTTTTGTGTGGATTGTTGAAACAGGGTCTCCCTTGGTCGCC 360
DB      301  TGCCTGGCTGATCTTTTGTGTGGATTGTTGAAACAGGGTCTCCCTTGGTCGCC 360

QY      361  AGGCTGGAGTCAGTGGTGCGATCTTGGCTCACTATAACCTCCACCTCCTGTTTCAAGT 420
DB      361  AGGCTGGAGTCAGTGGTGCGATCTTGGCTCACTATAACCTCCACCTCCTGTTTCAAGT 420

QY      421  GATCCTCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGACCCACACCCCG 480
DB      421  GATCCTCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGACCCACACCCCG 480

QY      481  CTAATTTTGTATTTTTATTAGAGACAGGGTTTCAACCATGTTGGCCAGGCTGTTCTCAA 540
DB      481  CTAATTTTGTATTTTTATTAGAGACAGGGTTTCAACCATGTTGGCCAGGCTGTTCTCAA 540

QY      541  CTCCTGGACTCAAGGGATCCGCTGCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG 600
DB      541  CTCCTGGACTCAAGGGATCCGCTGCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG 600

QY      601  AGTCACCATGCCTGACCTTATAATTCTTAAGTCAATTTTCTGGTCCATTTCTTCCTTAG 660
DB      601  AGTCACCATGCCTGACCTTATAATTCTTAAGTCAATTTTCTGGTCCATTTCTTCCTTAG 660

QY      661  GGTCTCACAACAATCTGCATTAGGGGTACAATAATCCTTAACCTTCACTGATTCAAAA 720
DB      661  GGTCTCACAACAATCTGCATTAGGGGTACAATAATCCTTAACCTTCACTGATTCAAAA 720

QY      721  AGGAAGATGAAGTATTGATTTAGAAAGGGGAAGTAGTACCCACTGCACACTCCT 780
DB      721  AGGAAGATGAAGTATTGATTTAGAAAGGGGAAGTAGTACCCACTGCACACTCCT 780
```

```

QY      781  GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAACAAA 840
DB      781  GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAACAAA 840

QY      841  TTTGGTTTAAATTAATTATCTAAATATCTAAATAATCTTTGGTATACATTTGTGATGTA 900
DB      841  TTTGGTTTAAATTAATTATCTAAATATCTAAATAATCTTTGGTATACATTTGTGATGTA 900

QY      901  ATGTAAGACTGTACAGACTTCTTAAATAATCTTTGGTCCATCTTTTCAATTTCCCCAG 960
DB      901  ATGTAAGACTGTACAGACTTCTTAAATAATCTTTGGTCCATCTTTTCAATTTCCCCAG 960

QY      961  TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAAGTGGCAGAGAAATTGACAATTGATT 1020
DB      961  TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAAGTGGCAGAGAAATTGACAATTGATT 1020

QY      1021  CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
DB      1021  CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAAAATGAAGACCTTACTGATG 1080

QY      1081  AACTAAGCTTGAATAAAATTTCTCTGATACACTACAGATAAATCTCGGGAATCTTAAACCAA 1140
DB      1081  AACTAAGCTTGAATAAAATTTCTCTGATACACTACAGATAAATCTCGGGAATCTTAAACCAA 1140

QY      1141  TTATGATGATGGCAACCAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
DB      1141  TTATGATGATGGCAACCAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200

QY      1201  ACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATTGATTGGTCTGTTACAGTCAAGCAA 1260
DB      1201  ACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATTGATTGGTCTGTTACAGTCAAGCAA 1260

QY      1261  TTGAAGCGCTTCCCCAGATAAATATGGCCAAATAGAGAGTTTGTGCTAGAAATTTCAAGTGA 1320
DB      1261  TTGAAGCGCTTCCCCAGATAAATATGGCCAAATAGAGAGTTTGTGCTAGAAATTTCAAGTGA 1320

QY      1321  GATTTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGCACTGACTACTTTCAAATGG 1380
DB      1321  GATTTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGCACTGACTACTTTCAAATGG 1380

QY      1381  CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCAAAATTTGAACCTGT 1440
DB      1381  CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCAAAATTTGAACCTGT 1440

QY      1441  CACAAGTAATGTCAAAAAAGTAACAACTTCTTCAAAAAAGCTGTAGAACGTGGAGCAG 1500
DB      1441  CACAAGTAATGTCAAAAAAGTAACAACTTCTTCAAAAAAGCTGTAGAACGTGGAGCAG 1500

QY      1501  TACCACTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
DB      1501  TACCACTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560

QY      1561  TTTCCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
DB      1561  TTTCCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620

QY      1621  TTTCCGGTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680
DB      1621  TTTCCGGTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680

QY      1681  CTACTAAAGCCAGGTTTATATATGGAGAGAAATCCACCAAGATGAGAAATAGGTT 1740
DB      1681  CTACTAAAGCCAGGTTTATATATGGAGAGAAATCCACCAAGATGAGAAATAGGTT 1740

QY      1741  ACCGGAATTCATTGAGACAAACTAACAAACTAAACAGTCAATGCCCATTTTGAAGAGTCC 1800
DB      1741  ACCGGAATTCATTGAGACAAACTAACAAACTAAACAGTCAATGCCCATTTTGAAGAGTCC 1800

QY      1801  CAGTTAACTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAAGTTGTACCTT 1860
DB      1801  CAGTTAACTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAAGTTGTACCTT 1860

QY      1861  GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTGGTTGTGCTGGATCTA 1920
```

Db 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTA 1920
QY 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGGTGTGATGAAAGAGTTCTGAACTTATTATTACTGATTCATATA 2040
Db 1981 TCAAGGAACCTCTGGTGTGATGAAAGAGTTCTGAACTTATTATTACTGATTCATATA 2040
QY 2041 CCCTGAAGATATAAACCAGGATCAAGTCTTCTAGCTAAATTAAGAACTAAAGAGTATC 2100
Db 2041 CCCTGAAGATATAAACCAGGATCAAGTCTTCTAGCTAAATTAAGAACTAAAGAGTATC 2100
QY 2101 AAGAACCAGAGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACCAGAGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
QY 2161 GTATTAAACCAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACCAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
QY 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTAAGACACCAA 2340
Db 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTACATGAGTGTGTTTAGAATCCAGTGTGTAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGTGTGTTTAGAATCCAGTGTGTAAGAAATGACT 2400
QY 2401 TTCCACCTGCTTGTGAGTTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGAGTTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAAAATACCTTGCACCTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATACCTTGCACCTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
QY 2521 ATGAATGCATTTCCGTTAAAGGAAGAAATTTATCCATATTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCATTTCCGTTAAAGGAAGAAATTTATCCATATTAAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGTATTTCAAGGTGTTAAATGAAAGAAACAGATATATGCTATAAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTCAAGGTGTTAAATGAAAGAAACAGATATATGCTATAAAATATG 2640
QY 2641 TGAACCTTAGAAGACAGATAACCAACTCTTGTATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGACAGATAACCAACTCTTGTATAGTTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTATGATTATGAAATCACGG 2760
Db 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTATGATTATGAAATCACGG 2760
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAA 2820
QY 2821 AGAAAAATCCATTCATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAATCCATTCATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
QY 2881 TTACACAAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAGCTAACTTCTGA 2940
Db 2881 TTACACAAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAGCTAACTTCTGA 2940
QY 2941 TAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTCGAAACCAAAATGCAACCCAGATA 3000

Db 2941 TAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTCGAAACCAAATGCAACCCAGATA 3000
QY 3001 CAACAAGTGTGTTTAAAGATTTCTCAGGTGGCACAGTTAATTATATGCCCAGAGCA 3060
Db 3001 CAACAAGTGTGTTTAAAGATTTCTCAGGTGGCACAGTTAATTATATGCCCAGAGCA 3060
QY 3061 TCAAAGATATGTTCTCTCCAGAGAGAAATGGAATCTAAGTCAAAAGATAAGCCCAAAA 3120
Db 3061 TCAAAGATATGTTCTCTCCAGAGAGAAATGGAATCTAAGTCAAAAGATAAGCCCAAAA 3120
QY 3121 GTGATGTTTGGTCTTAGGATGTTTGTACTATATGACTTACGGGAAAACACCATTTT 3180
Db 3121 GTGATGTTTGGTCTTAGGATGTTTGTACTATATGACTTACGGGAAAACACCATTTT 3180
QY 3181 ACCAGATAATTATCAGATTTCTAAATTACATGCCATAATTGATCCTAATCATGAATTT 3240
Db 3181 ACCAGATAATTATCAGATTTCTAAATTACATGCCATAATTGATCCTAATCATGAATTT 3240
QY 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGG 3300
Db 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGG 3300
QY 3301 ACCCAAAAACAGAGATATCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAATTTCAA 3360
Db 3301 ACCCAAAAACAGAGATATCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAATTTCAA 3360
QY 3361 CTCATCCAGTTAACCAAAATGGCCAAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAACCAAAATGGCCAAAGGAAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
QY 3421 AACTTGTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAACCTTTATATGAAC 3480
Db 3421 AACTTGTGGTCTGAATTTCTCTAACTCCATTTTGAAGCTGCTAAACCTTTATATGAAC 3480
QY 3481 ACTATAGTGGTGGTAAAGTCAATAATTTCTCATCTTCAAGACTTTTGAAGAAAGGG 3540
Db 3481 ACTATAGTGGTGGTAAAGTCAATAATTTCTCATCTTCAAGACTTTTGAAGAAAGGG 3540
QY 3541 GAAAAAAATGATTTGCAAGTATTTCGTAATGTCAGATAGGAGGTATAAATATATTTGGACT 3600
Db 3541 GAAAAAAATGATTTGCAAGTATTTCGTAATGTCAGATAGGAGGTATAAATATATTTGGACT 3600
QY 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACTCACTCTGAAGTGT 3660
Db 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACTCACTCTGAAGTGT 3660
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAACCTGTAAGAAATAGCAACCAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAACCTGTAAGAAATAGCAACCAC 3720
QY 3721 TTATGGCACTGTATATATTTGTAGACTTTGTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
Db 3721 TTATGGCACTGTATATATTTGTAGACTTTGTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGGTGGATAGCAAGTATATTTCTAAAAAATTTG 3840
Db 3781 TTGACATCATTTTACTCTTGGAAATAGTGGGTGGATAGCAAGTATATTTCTAAAAAATTTG 3840
QY 3841 TAAATAAAGTTTGTGGCTAAATGA 3866
Db 3841 TAAATAAAGTTTGTGGCTAAATGA 3866

RESULT 7
US-10-353-690-1
; Sequence 1, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.


```

; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodrique-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MPI02-018P1RNOMNIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3866
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-353-690-1

```

```

Query Match      100.0%; Score 3866; DB 16; Length 3866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGAATTCCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60
Db      1  GGAATTCCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60

QY      61  CAATGGCACAATCTCAGCTTACTGTCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGC 120
Db      61  CAATGGCACAATCTCAGCTTACTGTCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGC 120

QY      121  CTCAGCTCTCAAGTAGCTGGGATTACAGGCATGTGCCACCACCCCTGGCTAACTATT 180
Db      121  CTCAGCTCTCAAGTAGCTGGGATTACAGGCATGTGCCACCACCCCTGGCTAACTATT 180

QY      181  CTTTCTATTAGTAGAGATGGGTTTTCACCATGTTGGTCAGGCTGGTCTTGAACCTCG 240
Db      181  CTTTCTATTAGTAGAGATGGGTTTTCACCATGTTGGTCAGGCTGGTCTTGAACCTCG 240

QY      241  ACCTCAGGTGATCCACCTTGCCTTGGCTCCCAAGTGTCTAGGATTACAGCCGTGAAC 300
Db      241  ACCTCAGGTGATCCACCTTGCCTTGGCTCCCAAGTGTCTAGGATTACAGCCGTGAAC 300

QY      301  TGCTGGCTGATCTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 360
Db      301  TGCTGGCTGATCTCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 360

```

```

QY      361  AGGCTGGAGTGCAGTGGTGGATCTTGGCTCACTATAAACCCTCCACCTCCTGGTTCAAGT 420
Db      361  AGGCTGGAGTGCAGTGGTGGATCTTGGCTCACTATAAACCCTCCACCTCCTGGTTCAAGT 420

QY      421  GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGCGGTGCACACACACCCGG 480
Db      421  GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGCGGTGCACACACACCCGG 480

QY      481  CTAATTTTGTATTTTATTAGACAGGTTTACCATTGTTGGCCAGGCTGTTCTCAAA 540
Db      481  CTAATTTTGTATTTTATTAGACAGGTTTACCATTGTTGGCCAGGCTGTTCTCAAA 540

QY      541  CTCCTGGACTCAAGGGATCCGCTCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG 600
Db      541  CTCCTGGACTCAAGGGATCCGCTCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG 600

QY      601  AGTCACCATGCTGACCTTATAAATCTTAAGTCAATTTTCTGGTCCATTTCTTCTTAG 660
Db      601  AGTCACCATGCTGACCTTATAAATCTTAAGTCAATTTTCTGGTCCATTTCTTCTTAG 660

QY      661  GGTCTCAACAACAATCTGCATTAGGCGGTACAATAATCCTTAACCTCATGATTACAAA 720
Db      661  GGTCTCAACAACAATCTGCATTAGGCGGTACAATAATCCTTAACCTCATGATTACAAA 720

QY      721  AGGAAGATGAAGTGATTTCATGATTAGAAAGGGAGTAGTAAGCCCACTGCACACTCCT 780
Db      721  AGGAAGATGAAGTGATTTCATGATTAGAAAGGGAGTAGTAAGCCCACTGCACACTCCT 780

QY      781  GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAAAA 840
Db      781  GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAAAA 840

QY      841  TTTGGTTAAATTAATTAATAATATCTAAATAATCTTAAATAATCTTGGATACATTG 900
Db      841  TTTGGTTAAATTAATTAATAATATCTAAATAATCTTAAATAATCTTGGATACATTG 900

QY      901  ATGTAAGACTGTACAGACTTCTAGAAAACAGTTTGGGTTCCATCTTTTCAATTTCCC 960
Db      901  ATGTAAGACTGTACAGACTTCTAGAAAACAGTTTGGGTTCCATCTTTTCAATTTCCC 960

QY      961  TGCAGTTTCTGTAGAAATGGAATCCGAGGATTAAGTGGCAGAGAAATTGACAAATTG 1020
Db      961  TGCAGTTTCTGTAGAAATGGAATCCGAGGATTAAGTGGCAGAGAAATTGACAAATTG 1020

QY      1021  CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAAAATGAAGACCTTACTG 1080
Db      1021  CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAAAATGAAGACCTTACTG 1080

QY      1081  AACTAAGCTTGAATAAAATTTCTGTGATACAGATAAATCTCGGAACTGTTAACCAA 1140
Db      1081  AACTAAGCTTGAATAAAATTTCTGTGATACAGATAAATCTCGGAACTGTTAACCAA 1140

QY      1141  TTATGATGATGGCAACAACCCAGAGGACTGGTTGAGTTTGTCTCAAAACTAGAGAAA 1200
Db      1141  TTATGATGATGGCAACAACCCAGAGGACTGGTTGAGTTTGTCTCAAAACTAGAGAAA 1200

QY      1201  ACAGTGTCCGCTAAGTAGTGTCTTTTAAATAAATTTGATTGGTGGTCAAGCAA 1260
Db      1201  ACAGTGTCCGCTAAGTAGTGTCTTTTAAATAAATTTGATTGGTGGTCAAGCAA 1260

QY      1261  TTGAAGCGCTTCCCCAGATATAATATGGCCAAAATGAGAGTTTGTCTAGAAATCA 1320
Db      1261  TTGAAGCGCTTCCCCAGATATAATATGGCCAAAATGAGAGTTTGTCTAGAAATCA 1320

QY      1321  GATTGTCTGAATTAAGCTATTCAAGAGCCAGATGATGCAGTGACTACTTTCAATGG 1380
Db      1321  GATTGTCTGAATTAAGCTATTCAAGAGCCAGATGATGCAGTGACTACTTTCAATGG 1380

QY      1381  CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTG 1440
Db      1381  CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTG 1440

QY      1441  CACAAGGTAATGTCAAAAAAAGTAACAACTTCTTCAAAAAAGTGTAGAACGGAGC 1500

```

Db 1441 CACAAGGTAATGTCAAAAAAGTAAACAACACTTCTTCAAAAAAGCTGTAGACGTGGAGCAG 1500
Qy 1501 TACCAC TAGAATAGCTGGAAATTCGCCTGCGGAATTTAAACCTCCAAAAAGACAGCTGC 1560
Db 1501 TACCAC TAGAATAGCTGGAAATTCGCCTGCGGAATTTAAACCTCCAAAAAGACAGCTGC 1560
Qy 1561 TTTTCAGAGGAGAAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
Db 1561 TTTTCAGAGGAGAAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
Qy 1621 TTTTCGGTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680
Db 1621 TTTTCGGTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680
Qy 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACACAAGATGCAGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACACAAGATGCAGAAATAGGTT 1740
Qy 1741 ACCGGAATTCATGAGACAAAACCTAACAAAACTAAACAGTCATGCCCATTTGGAAGAGTCC 1800
Db 1741 ACCGGAATTCATGAGACAAAACCTAACAAAACTAAACAGTCATGCCCATTTGGAAGAGTCC 1800
Qy 1801 CAGTTAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Qy 1861 GTTTTATGAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCTGGATCTA 1920
Db 1861 GTTTTATGAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCTGGATCTA 1920
Qy 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Qy 1981 TCAAGGAACCTCTGGTGTGAGTGAAGAGATTCTGAACCTTATTATTACTGATCAATAA 2040
Db 1981 TCAAGGAACCTCTGGTGTGAGTGAAGAGATTCTGAACCTTATTATTACTGATCAATAA 2040
Qy 2041 CCCTGAAGAAATAAAACGGAATCAAGTCTTCTAGCTAAATTAAGAGAAACCTAAAGAGTATC 2100
Db 2041 CCCTGAAGAAATAAAACGGAATCAAGTCTTCTAGCTAAATTAAGAGAAACCTAAAGAGTATC 2100
Qy 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGTAAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACAGAAACAGTGGCAAGTAAAGAGAAAGTCAGAGT 2160
Qy 2161 GTATTAAACCCAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCGAA 2220
Db 2161 GTATTAAACCCAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCGAA 2220
Qy 2221 AAGTTAATACAGAGCAGAAACATACCACITTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACITTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAAC 2280
Qy 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA 2340
Db 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
Qy 2401 TTCCACCTGCTTGTGATGTCAACACCTTATGGCCAAACCTGCCTTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGATGTCAACACCTTATGGCCAAACCTGCCTTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAAATACCTTGCCACTCCACTTCAAAATTTACAGGTTTATAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATACCTTGCCACTCCACTTCAAAATTTACAGGTTTATAGCATCTTCTTCAGCAA 2520
Qy 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATCCATATTAAGCAGATAGGAAGTGGAG 2580

Db 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATCCATATTTAAAGCAGATAGGAAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGTATTTCAAGGTGTAAATGAAAGAAACAGATATATGCTATAAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTCAAGGTGTAAATGAAAGAAACAGATATATGCTATAAAATATG 2640
Qy 2641 TGAACCTTAGAAGAACGAGATAACCAAACTCTTGATAGTTACGGAAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAACGAGATAACCAAACTCTTGATAGTTACGGAAACGAAATAGCTTATT 2700
Qy 2701 TGAATAAACTACAACAACACAGTGATAAGATCATCCGACTTTATGATTATGAAATCAGCG 2760
Db 2701 TGAATAAACTACAACAACACAGTGATAAGATCATCCGACTTTATGATTATGAAATCAGCG 2760
Qy 2761 ACCAGTACATCTACATGGTAATGGAGTGTGAAATATTGATCTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGAAATATTGATCTTAATAGTTGGCTTAAAA 2820
Qy 2821 AGAAAAAATCCATTCATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGSCAG 2880
Db 2821 AGAAAAAATCCATTCATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGSCAG 2880
Qy 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Db 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Qy 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTTGGGATTGCAAAACCAAAATGCAACCAAGTA 3000
Db 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTTGGGATTGCAAAACCAAAATGCAACCAAGTA 3000
Qy 3001 CAAACAAGTGTGTTAAAGATTTCTCAGGTTGGCACAGTTAATTAATATGCCACCAGAAAGCAA 3060
Db 3001 CAAACAAGTGTGTTAAAGATTTCTCAGGTTGGCACAGTTAATTAATATGCCACCAGAAAGCAA 3060
Qy 3061 TCAAAAGATATGCTTCTCCAGAGAGAAATGGGAAAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
Db 3061 TCAAAAGATATGCTTCTCCAGAGAGAAATGGGAAAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
Qy 3121 GTGATGTTTGGTCTTAGGATGTTTGTGACTATATGACTTACGGGAAAAACCACTATTC 3180
Db 3121 GTGATGTTTGGTCTTAGGATGTTTGTGACTATATGACTTACGGGAAAAACCACTATTC 3180
Qy 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCCTAAATCATGAAATTG 3240
Db 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCCTAAATCATGAAATTG 3240
Qy 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Qy 3301 ACCCAAAACAGAGGATATCCATTCCTGAGCTCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db 3301 ACCCAAAACAGAGGATATCCATTCCTGAGCTCTGGCTCATCCATATGTTCAAATTCAAA 3360
Qy 3361 CTATCCAGTTAACCAAAATGGCCAAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
Db 3361 CTATCCAGTTAACCAAAATGGCCAAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
Qy 3421 AACTTGTGTTGTTCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
Db 3421 AACTTGTGTTGTTCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAAC 3480
Qy 3481 ACTATAGTGGTGGTGAAGAGTCATAATTTCTTCATCTCCAGACTTTTGAAAAAAAAAGGG 3540
Db 3481 ACTATAGTGGTGGTGAAGAGTCATAATTTCTTCATCTCCAGACTTTTGAAAAAAAAAGGG 3540
Qy 3541 GAAAAAATGATTTGCAGTTATTCGTAATGTCAGATAGGAGGTATAAAATATATTTGGACT 3600
Db 3541 GAAAAAATGATTTGCAGTTATTCGTAATGTCAGATAGGAGGTATAAAATATATTTGGACT 3600
Qy 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db 3601 GTTATACTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660

QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAAAAAGTGAATAATAGCAACCAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAAAAAGTGAATAATAGCAACCAC 3720
QY 3721 TTATGGCACTGATATATTTAGAGCTTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
Db 3721 TTATGGCACTGATATATTTAGAGCTTTTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
QY 3781 TTGACATCATTTTACTCTTGGATAGTGGTGGATAGCAAGTATATTTCTAAAAAATTTG 3840
Db 3781 TTGACATCATTTTACTCTTGGATAGTGGTGGATAGCAAGTATATTTCTAAAAAATTTG 3840
QY 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866
Db 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866

RESULT 8
US-10-366-288-15
; Sequence 15, Application US/10366288
; Publication No. US20030216288A1
; GENERAL INFORMATION:
; APPLICANT: Powell, Douglas
; APPLICANT: Weich, Nadine S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; AIDS AND HIV-RELATED DISORDERS USING 1414, 1481, 1553,
; TITLE OF INVENTION: 34021, 1720, 1683, 1552, 1682, 1675, 12825, 9952, 5816,
; TITLE OF INVENTION: 10002, 1611, 1371, 14324, 126, 270, 312, 167, 326, 18926,
; TITLE OF INVENTION: 6747, 1793, 1784 OR 2045 MOLECULES
; FILE REFERENCE: MPI02-025P1RNMNIM
; CURRENT APPLICATION NUMBER: US/10/366,288
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 60/357,391
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/380,249
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 60/391,306
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/406,297
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 60/412,007
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 60/417,508
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 60/432,318
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3866
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-366-288-15

Query Match 100.0%; Score 3866; DB 16; Length 3866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAATTCCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60
Db 1 GGAATTCCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60
QY 61 CAATGGCACAACTCAGCTTACTGCAACCTCCGCCCTCCCGGTTCAAGCGATTCTCCTGC 120
Db 61 CAATGGCACAACTCAGCTTACTGCAACCTCCGCCCTCCCGGTTCAAGCGATTCTCCTGC 120
QY 121 CTCAGCCTCTCAAGTAGTGGGATTACAGGATGTGCCACCCCTGGCTAACTAATTT 180
Db 121 CTCAGCCTCTCAAGTAGTGGGATTACAGGATGTGCCACCCCTGGCTAACTAATTT 180
QY 181 CTTTCTATTTAGTAGAGATGGGTTTCAACATGTTGGTCAGGCTGGTCTTGAACCTCCTG 240
Db 181 CTTTCTATTTAGTAGAGATGGGTTTCAACATGTTGGTCAGGCTGGTCTTGAACCTCCTG 240

Db 181 CTTTCTATTTAGTAGAGATGGGTTTCAACATGTTGGTCAGGCTGGTCTTGAACCTCCTG 240
QY 241 ACCTCAGGTGATCCACTTGCCTTGGCTCCCAAAGTGTAGGATTACAGCCGTGAAACTG 300
Db 241 ACCTCAGGTGATCCACTTGCCTTGGCTCCCAAAGTGTAGGATTACAGCCGTGAAACTG 300
QY 301 TGCCTGGCTGATTTCTTTTGTGTTGGATTTTGAACACAGGGTCTCCCTTGGTCGCC 360
Db 301 TGCCTGGCTGATTTCTTTTGTGTTGGATTTTGAACACAGGGTCTCCCTTGGTCGCC 360
QY 361 AGGCTGGAGTGCAGTGGTGGATCTTGGCTCACTATAAACCCTCCACCTCTGGTTCAAGT 420
Db 361 AGGCTGGAGTGCAGTGGTGGATCTTGGCTCACTATAAACCCTCCACCTCTGGTTCAAGT 420
QY 421 GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGCACCCACACCCGG 480
Db 421 GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGCACCCACACCCGG 480
QY 481 CTAATTTTGTATTTTATTAGACAGGGTTTCAACATGTTGGCCAGGCTGTTCTCAAA 540
Db 481 CTAATTTTGTATTTTATTAGACAGGGTTTCAACATGTTGGCCAGGCTGTTCTCAAA 540
QY 541 CTCCTGGACTCAAGGGATCCGCTGCCCTCCACCTCCCAAAGTCCCGAGATTACAGGTGTG 600
Db 541 CTCCTGGACTCAAGGGATCCGCTGCCCTCCACCTCCCAAAGTCCCGAGATTACAGGTGTG 600
QY 601 AGTCACCATGCTGACCTTATAATCTTAAGTCAATTTTCTGGTCCATTTCTTCTTAG 660
Db 601 AGTCACCATGCTGACCTTATAATCTTAAGTCAATTTTCTGGTCCATTTCTTCTTAG 660
QY 661 GGTCTCACAACAAATCTGATTAGGCGGTACAATAATCTTAACTTCAATGATTACAAA 720
Db 661 GGTCTCACAACAAATCTGATTAGGCGGTACAATAATCTTAACTTCAATGATTACAAA 720
QY 721 AGGAAGATGAAGTGATTTCATGATTAGAAAGGGGAAGTAGTAAGCCACTGCACATCCT 780
Db 721 AGGAAGATGAAGTGATTTCATGATTAGAAAGGGGAAGTAGTAAGCCACTGCACATCCT 780
QY 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGATGGGAAGGTAGAATACAAA 840
Db 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGATGGGAAGGTAGAATACAAA 840
QY 841 TTTGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
Db 841 TTTGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
QY 901 ATGTAAGACTGTACAGACTTCTAGAAACAGTGGTGGTCCATCTTTTCAATTTCCCCAG 960
Db 901 ATGTAAGACTGTACAGACTTCTAGAAACAGTGGTGGTCCATCTTTTCAATTTCCCCAG 960
QY 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATGACAATTGATT 1020
Db 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATGACAATTGATT 1020
QY 1021 CCATAATGAACAAAGTGAGAGACATTTAAATAAGTTTAAATAAGTTTAAATAAGTTTAA 1080
Db 1021 CCATAATGAACAAAGTGAGAGACATTTAAATAAGTTTAAATAAGTTTAAATAAGTTTAA 1080
QY 1081 AACTAAGCTTGAATAAATTTCTGCTGATACAGATAAATCTCGGAACTGTTAACCCAAA 1140
Db 1081 AACTAAGCTTGAATAAATTTCTGCTGATACAGATAAATCTCGGAACTGTTAACCCAAA 1140
QY 1141 TTATGATGATGGCAACAAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
Db 1141 TTATGATGATGGCAACAAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
QY 1201 ACAGTGTTCGGCTAAGTAGTCTCTTTTAAATAAATTTGATTGGTCGTACAGTCAAGCAA 1260
Db 1201 ACAGTGTTCGGCTAAGTAGTCTCTTTTAAATAAATTTGATTGGTCGTACAGTCAAGCAA 1260
QY 1261 TTGAAGCGCTTCCCCAGATAAATATGGCCAAAATGAGAGTTTGTGCTAGAAATTCAGTGA 1320
Db 1261 TTGAAGCGCTTCCCCAGATAAATATGGCCAAAATGAGAGTTTGTGCTAGAAATTCAGTGA 1320

Qy 1321 GATTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG 1380
Db 1321 GATTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGG 1380
Qy 1381 CCAGAGCAAACTGCAAGAAATTTGGCTTTTGTTCATATATCTTTGTCACAAATTTGAACTGT 1440
Db 1381 CCAGAGCAAACTGCAAGAAATTTGGCTTTTGTTCATATATCTTTGTCACAAATTTGAACTGT 1440
Qy 1441 CACAAGGTAATGTCAAAAAAGTAACAACACTTCTTCAAAAAGCTGTAGAACGTTGGAGCAG 1500
Db 1441 CACAAGGTAATGTCAAAAAAGTAACAACACTTCTTCAAAAAGCTGTAGAACGTTGGAGCAG 1500
Qy 1501 TACCAC TAGAAATGCTGGAAATTTGCCCTGCGGAAATTTAACTCCAAAAAAGCAGCTGC 1560
Db 1501 TACCAC TAGAAATGCTGGAAATTTGCCCTGCGGAAATTTAACTCCAAAAAAGCAGCTGC 1560
Qy 1561 TTTTCAGAGGAGGAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
Db 1561 TTTTCAGAGGAGGAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
Qy 1621 TTTCCGGTTCACTTGGGCATTTACAGAAATPAGGAACAACAGTTGTGATTTCCAGAGGACAGA 1680
Db 1621 TTTCCGGTTCACTTGGGCATTTACAGAAATPAGGAACAACAGTTGTGATTTCCAGAGGACAGA 1680
Qy 1681 CTACTAAAGCCAGTTTATATGGAGAGAAACATGCCACCACAAGATGCAGAAATPAGGTT 1740
Db 1681 CTACTAAAGCCAGTTTATATGGAGAGAAACATGCCACCACAAGATGCAGAAATPAGGTT 1740
Qy 1741 ACCGGAATTCATTGAGACAAAACCTAACAACACTAAACAGTCATGCCCATTTGGAAGAGTCC 1800
Db 1741 ACCGGAATTCATTGAGACAAAACCTAACAACACTAAACAGTCATGCCCATTTGGAAGAGTCC 1800
Qy 1801 CAGTTAAACCTTCTAAATAGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAAACCTTCTAAATAGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Qy 1861 GTTTTATGAAAAGCAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGAATCTA 1920
Db 1861 GTTTTATGAAAAGCAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGAATCTA 1920
Qy 1921 AACCAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 AACCAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Qy 1981 TCAAGGAACCTCTGGTGTGATGAAAAGAGTCTGAACTTATATTAATGATTCATATAA 2040
Db 1981 TCAAGGAACCTCTGGTGTGATGAAAAGAGTCTGAACTTATATTAATGATTCATATAA 2040
Qy 2041 CCTGTAAGAAATAAAACCGAATCAAGTCTTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
Db 2041 CCTGTAAGAAATAAAACCGAATCAAGTCTTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
Qy 2101 AAGAACAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Qy 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCGAA 2220
Db 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCGAA 2220
Qy 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAAAAC 2280
Qy 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTAAGACACCAA 2340
Db 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTGTAAGACACCAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAGAACTCCAGTTGTAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAGAACTCCAGTTGTAAGAAATGACT 2400

Qy 2401 TTCCACCTGCTTGTGACGTTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGACGTTGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAATACCTTGCACCTCCACTTCAAAATTTACAGTTTTCAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAATACCTTGCACCTCCACTTCAAAATTTACAGTTTTCAGCATCTTCTTCAGCAA 2520
Qy 2521 ATGAATGCATTTCCGTTAAAGGAAGAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCATTTCCGTTAAAGGAAGAATTTATTTCCATATTTAAAGCAGATAGGAAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGTAATTTCAAGGTGTTAAATGAAAAAGAAACAGATATATCTATAAAATATG 2640
Db 2581 GTTCAAGCAAGGTAATTTCAAGGTGTTAAATGAAAAAGAAACAGATATATCTATAAAATATG 2640
Qy 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGATAGTTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAGCAGATAACCAACTCTTGATAGTTTACCGGAACGAAATAGCTTATT 2700
Qy 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCACGG 2760
Db 2701 TGAATAAACTACAACACACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCACGG 2760
Qy 2761 ACCAGTACATCTACATGTAATGGAGTGTGAAATATTTGATCTTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGTAATGGAGTGTGAAATATTTGATCTTTAATAGTTGGCTTAAAA 2820
Qy 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Qy 2881 TTCAACAAATCCATCAACATGGCATTTGTTCAAGTGATCTTTAAACCCAGCTAACTTTCTGA 2940
Db 2881 TTCAACAAATCCATCAACATGGCATTTGTTCAAGTGATCTTTAAACCCAGCTAACTTTCTGA 2940
Qy 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTTGGGATGCAAAACCAATGCAACCCAGATA 3000
Db 2941 TAGTTGATGGAATGCTAAAGCTAATTTGATTTTGGGATGCAAAACCAATGCAACCCAGATA 3000
Qy 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCACAGTTAATTTATATGCCACCAGAAAGCAA 3060
Db 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCACAGTTAATTTATATGCCACCAGAAAGCAA 3060
Qy 3061 TCAAGATATGCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
Db 3061 TCAAGATATGCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
Qy 3121 GTGATGTTTGGTCTTAGGATGTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCTTAGGATGTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Qy 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCCTTAATCATGAAATTG 3240
Db 3181 AGCAGATAATTAATCAGATTTCTAAATTTACATGCCATAATTTGATCCTTAATCATGAAATTG 3240
Qy 3241 AATTTCCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCCGATATTTCCAGAGAAAGATCTTCAAGATGTTTAAAGTGTGTTTAAAAAGGG 3300
Qy 3301 ACCCAAAACAGAGGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db 3301 ACCCAAAACAGAGGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Qy 3361 CTCACTCAGTTAACCAATGGCCAAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
Db 3361 CTCACTCAGTTAACCAATGGCCAAAGGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
Qy 3421 AACTTGTGTTGTTCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTAAACTTTTATATGAAC 3480
Db 3421 AACTTGTGTTGTTCTGAATTTCTCCTAACTCCATTTTGAAGCTGCTAAACTTTTATATGAAC 3480
Qy 3481 ACTATAGTGGTGGTGAAAGTCATAATTTCTTCACTCTCCAAGACTTTTGGAAAAAAGGG 3540

Db 961 |||||TGCGATTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAATTGATT 1020
Qy 1021 CCATAATGAACAAAGTGAGAGACATTAAAAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
Db 1021 |||||CCATAATGAACAAAGTGAGAGACATTAAAAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
Qy 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACACAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1140
Db 1081 |||||AACTAAGCTTGAATAAAATTTCTGCTGATACACAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1140
Qy 1141 TTATGATGATGSCAAAACAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
Db 1141 |||||TTATGATGATGSCAAAACAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
Qy 1201 ACAGTGTTCGGCTAAAGTGATGCTCTTTTAAATAAATTTGATGGTGGTTACAGTCAAGCAA 1260
Db 1201 |||||ACAGTGTTCGGCTAAAGTGATGCTCTTTTAAATAAATTTGATGGTGGTTACAGTCAAGCAA 1260
Qy 1261 TTGAAGCGCTTCCCCCAGATAAAATATGCCCAAATGAGAGTTTGTCTAGAAATCAAGTGA 1320
Db 1261 |||||TTGAAGCGCTTCCCCCAGATAAAATATGCCCAAATGAGAGTTTGTCTAGAAATCAAGTGA 1320
Qy 1321 GATTTGCTGAATTAAGAGCTATTTCAAGAGCCAGATGATGACGTGACTACTTTCAAATGG 1380
Db 1321 |||||GATTTGCTGAATTAAGAGCTATTTCAAGAGCCAGATGATGACGTGACTACTTTCAAATGG 1380
Qy 1381 CCAGAGCAAACCTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTGAACTGT 1440
Db 1381 |||||CCAGAGCAAACCTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTGAACTGT 1440
Qy 1441 CACAAGGTAATGTCAAAAAAGTAAACAACCTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 |||||CACAAGGTAATGTCAAAAAAGTAAACAACCTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Qy 1501 TACCAC TAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
Db 1501 |||||TACCAC TAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
Qy 1561 TTTCAGAGGAGGAAAAAGAAATTTATCAGCATCTACGGTATTAACCTGCCCCAAGAATCAT 1620
Db 1561 |||||TTTCAGAGGAGGAAAAAGAAATTTATCAGCATCTACGGTATTAACCTGCCCCAAGAATCAT 1620
Qy 1621 TTTCCGGTTACCTTGGGCATTTACAGAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680
Db 1621 |||||TTTCCGGTTACCTTGGGCATTTACAGAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680
Qy 1681 CTACTAAAGCCAGGTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTT 1740
Db 1681 |||||CTACTAAAGCCAGGTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTT 1740
Qy 1741 ACCGGAATTCATTGAGACAACTAACAAAACTAAACAGTCATGCCCATTTTGGAGAGTCC 1800
Db 1741 |||||ACCGGAATTCATTGAGACAACTAACAAAACTAAACAGTCATGCCCATTTTGGAGAGTCC 1800
Qy 1801 CAGTTAACCTTCTAAATAGCCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 |||||CAGTTAACCTTCTAAATAGCCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Qy 1861 GTTTTATGAAAAGCAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCCTGGATCTA 1920
Db 1861 |||||GTTTATGAAAAGCAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCCTGGATCTA 1920
Qy 1921 AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Db 1921 |||||AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT 1980
Qy 1981 TCAAGGAACCTCTGGTGTGAGATGAAAAGAGTTCTGAACCTTATTTACTGATTCATATAA 2040
Db 1981 |||||TCAAGGAACCTCTGGTGTGAGATGAAAAGAGTTCTGAACCTTATTTACTGATTCATATAA 2040
Qy 2041 CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
Db 2041 |||||CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC 2100

Db 2041 CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
Qy 2101 AAGAAC CAGAGGTTCCAGAGAGTAACCGAAGACAGTGGCAAGCTTAAGAGAAAAGTCAGAGT 2160
Db 2101 |||||AAGAAC CAGAGGTTCCAGAGAGTAACCGAAGACAGTGGCAAGCTTAAGAGAAAAGTCAGAGT 2160
Qy 2161 GTATTAAACCAGAAATCCTGCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Db 2161 |||||GTATTAAACCAGAAATCCTGCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Qy 2221 AAGTTAATACAGAGCAGAAAAACATACCCTTTTGAGCAA CCTGTCTTTTCAGTTTCAAAAAC 2280
Db 2221 |||||AAGTTAATACAGAGCAGAAAAACATACCCTTTTGAGCAA CCTGTCTTTTCAGTTTCAAAAAC 2280
Qy 2281 AGTCACCAATATCAACATCTAAATGGTTTGACCCCAAATCTATTTGTAAGACACCAA 2340
Db 2281 |||||AGTCACCAATATCAACATCTAAATGGTTTGACCCCAAATCTATTTGTAAGACACCAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAATGACT 2400
Db 2341 |||||GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAATGACT 2400
Qy 2401 TTCCACCTGCTTGTGATGTTGTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
Db 2401 |||||TTCCACCTGCTTGTGATGTTGTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAAATACCTTGCACCTCCACTTCAAATTTACAGGTTTTAGCATCTTCTTCAGCAA 2520
Db 2461 |||||AGCATCAAAATACCTTGCACCTCCACTTCAAATTTACAGGTTTTAGCATCTTCTTCAGCAA 2520
Qy 2521 ATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
Db 2521 |||||ATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGTAATTTGAGGTTTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
Db 2581 |||||GTTCAAGCAAGGTAATTTGAGGTTTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
Qy 2641 TGAACCTTAGAAGAACAGATAACCAAACCTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 |||||TGAACCTTAGAAGAACAGATAACCAAACCTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Qy 2701 TGAATAAACTACAACACACAGTGATAGATCATCCGACTTTTATGATTAATAATCACGG 2760
Db 2701 |||||TGAATAAACTACAACACACAGTGATAGATCATCCGACTTTTATGATTAATAATCACGG 2760
Qy 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATAATTGATCTTAAATAGTTGGCTTAAAA 2820
Db 2761 |||||ACCAGTACATCTACATGGTAATGGAGTGTGGAATAATTGATCTTAAATAGTTGGCTTAAAA 2820
Qy 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 |||||AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Qy 2881 TTCAACAATCCATCAACATGGCATTTGTTCAAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Db 2881 |||||TTCAACAATCCATCAACATGGCATTTGTTCAAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Qy 2941 TAGTTGATGGAATGCTAAAAGCTAATTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGATA 3000
Db 2941 |||||TAGTTGATGGAATGCTAAAAGCTAATTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGATA 3000
Qy 3001 CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGAGTTAAATATATGCCACCAGAAAGCAA 3060
Db 3001 |||||CAACAAGTGTGTTAAAGATTTCTCAGGTTGGCAGAGTTAAATATATGCCACCAGAAAGCAA 3060
Qy 3061 TCAAAAGATATGTTCTCCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCCAAAA 3120
Db 3061 |||||TCAAAAGATATGTTCTCCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCCAAAA 3120
Qy 3121 GTGATGTTTGGTCCCTTAGGATGTTATTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db 3121 |||||GTGATGTTTGGTCCCTTAGGATGTTATTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180

QY 661 GGTCTTCACACAAATCTGCATTAGCGGTACAATAATCCTTAACCTTCAATGATTCACAAA 720
Db 661 GGTCTTCACACAAATCTGCATTAGCGGTACAATAATCCTTAACCTTCAATGATTCACAAA 720
QY 721 AGGAAGATGAAGTGATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCCACTGCACACTCCT 780
Db 721 AGGAAGATGAAGTGATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCCACTGCACACTCCT 780
QY 781 GGATGATGATCCTAAATCCAGATACAGTAAAAATGGGGTATGGGAAGGTAGAAATACAAA 840
Db 781 GGATGATGATCCTAAATCCAGATACAGTAAAAATGGGGTATGGGAAGGTAGAAATACAAA 840
QY 841 TTTGGTTTAAATTAATTATCTAAATATCTAAAAACATTTTGGGATACATTTGATGTGA 900
Db 841 TTTGGTTTAAATTAATTATCTAAATATCTAAAAACATTTTGGGATACATTTGATGTGA 900
QY 901 ATGTAAGACTGTACAGACTTCTCTAGAAAAACAGTTTGGGTCCATCTTTTCATTTCCCCAG 960
Db 901 ATGTAAGACTGTACAGACTTCTCTAGAAAAACAGTTTGGGTCCATCTTTTCATTTCCCCAG 960
QY 961 TGCAGTTTCTGTAGAAAATGGAATCCGAGGATTTAAAGTGCAGAGAAATTGACAATTGATT 1020
Db 961 TGCAGTTTCTGTAGAAAATGGAATCCGAGGATTTAAAGTGCAGAGAAATTGACAATTGATT 1020
QY 1021 CCATAATGAACAAGTGAGAGACATTAATAAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
Db 1021 CCATAATGAACAAGTGAGAGACATTAATAAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
QY 1081 AACTAAGCTTGAATAAATTTCTGCTGATACAGATAAATCGGAACTGTTAACCAAA 1140
Db 1081 AACTAAGCTTGAATAAATTTCTGCTGATACAGATAAATCGGAACTGTTAACCAAA 1140
QY 1141 TTATGATGATGSCAAAACAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
Db 1141 TTATGATGATGSCAAAACAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
QY 1201 ACAGTGTTCGCTAAGTGATGCTCTTTTAAATAAATTTGATGGTGGTTCACAGTCAAGCAA 1260
Db 1201 ACAGTGTTCGCTAAGTGATGCTCTTTTAAATAAATTTGATGGTGGTTCACAGTCAAGCAA 1260
QY 1261 TTGAAGCGCTTCCCCAGATAAATATGCCCCAAAATGAGAGTTTGTCTAGAAATTCAGTGA 1320
Db 1261 TTGAAGCGCTTCCCCAGATAAATATGCCCCAAAATGAGAGTTTGTCTAGAAATTCAGTGA 1320
QY 1321 GATTTGCTGAATTTAAAAGCTATTCAAGAGCCAGATGATGACGCTGACTACTTTCAAATGG 1380
Db 1321 GATTTGCTGAATTTAAAAGCTATTCAAGAGCCAGATGATGACGCTGACTACTTTCAAATGG 1380
QY 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGCATATATCTTTTGCACAAATTTGAACTGT 1440
Db 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGCATATATCTTTTGCACAAATTTGAACTGT 1440
QY 1441 CACAAGGTAATGTCAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAATGTCAAAAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
QY 1501 TACCACCTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
Db 1501 TACCACCTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
QY 1561 TTTTCAGAGGAGGAAAGAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
Db 1561 TTTTCAGAGGAGGAAAGAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT 1620
QY 1621 TTTCCGGTTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTTGTGATTCAGAGGACAGA 1680
Db 1621 TTTCCGGTTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTTGTGATTCAGAGGACAGA 1680
QY 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCACAGATGCAGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCACAGATGCAGAAATAGGTT 1740

QY 1741 ACCGGAATTCATTGAGACAAAACCTAACAAAACCTAAACAGTCACTGCCCCATTTGGAAAGATCC 1800
Db 1741 ACCGGAATTCATTGAGACAAAACCTAAACAAAACCTAAACAGTCACTGCCCCATTTGGAAAGATCC 1800
QY 1801 CAGTTAACCTTCTAAATAGCCCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAACCTTCTAAATAGCCCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
QY 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCCTGGATCTA 1920
QY 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGCTGTGATGAAAAGAGTTCTGAACTTATTTACTGATTCATATAA 2040
Db 1981 TCAAGGAACCTCTGCTGTGATGAAAAGAGTTCTGAACTTATTTACTGATTCATATAA 2040
QY 2041 CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTTAGAAGAAAACCTAAAGAGTATC 2100
Db 2041 CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTTAGAAGAAAACCTAAAGAGTATC 2100
QY 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACGAGAAACAGTGGCAAGCTAAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACGAGAAACAGTGGCAAGCTAAAGAGAAAGTCAGAGT 2160
QY 2161 GTATTAAACCAGAAATCCTGCTGATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCCGAA 2220
Db 2161 GTATTAAACCAGAAATCCTGCTGATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAAAAACATACCCTTTTGGCAACCTGTCTTTTTCAGTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAAAAACATACCCTTTTGGCAACCTGTCTTTTTCAGTTTCAAAAC 2280
QY 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCCAAAATCTATTTGTAAGACACCAA 2340
Db 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCCAAAATCTATTTGTAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTTAGAACTCCAGTTGTAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTTAGAACTCCAGTTGTAAGAAATGACT 2400
QY 2401 TTCCACCTGCTGTGATGTTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTGTGATGTTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAAAATACCTTGCCACTCCACTTCAAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATACCTTGCCACTCCACTTCAAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
QY 2521 ATGAATGCATTTTCGGTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCATTTTCGGTTAAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATATAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATATAAATATG 2640
QY 2641 TGAACCTTAGAAGAACAGATAAACCAACTCTTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAACAGATAAACCAACTCTTTGATAGTTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAACACAGTGAAGATCATCCGACTTTTATGATTTATGAAATCACGG 2760
Db 2701 TGAATAAACTACAACACAGTGAAGATCATCCGACTTTTATGATTTATGAAATCACGG 2760
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATAATTTGATCTTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATAATTTGATCTTTAATAGTTGGCTTAAAA 2820
QY 2821 AGAAAAAATCCATTGATCCATGGGAACCGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880

Db 301 TGCCTGGCTGATCTCTTTTGGTTGGATTTTGGAAACAGGGTCTCCCTTGGTCGCC 360
Qy 361 AGGCTGGAGTGCAGTGGTGGGATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT 420
Db 361 AGGCTGGAGTGCAGTGGTGGGATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGT 420
Qy 421 GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGCAACCAACACCCGG 480
Db 421 GATCCTCCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGCAACCAACACCCGG 480
Qy 481 CTAATTTTGTATTTTATTAGACAGGGTTTACCAATGTTGGCCAGGCTGTTCTCAA 540
Db 481 CTAATTTTGTATTTTATTAGACAGGGTTTACCAATGTTGGCCAGGCTGTTCTCAA 540
Qy 541 CTCCTGGACTCAAGGGATCCGCTGCTCCACTTCCCAAGTCCCAGATTACAGGTGTG 600
Db 541 CTCCTGGACTCAAGGGATCCGCTGCTCCACTTCCCAAGTCCCAGATTACAGGTGTG 600
Qy 601 AGTCACCATGCTGACCTTATAATTTCTTAAGTCAATTTTCTGGTCCATTTCTTCCTTAG 660
Db 601 AGTCACCATGCTGACCTTATAATTTCTTAAGTCAATTTTCTGGTCCATTTCTTCCTTAG 660
Qy 661 GGTCCTCACACAAATCTGCATTTAGGCGGTACAATAATCCTTAACCTTCAATGATTCACAAA 720
Db 661 GGTCCTCACACAAATCTGCATTTAGGCGGTACAATAATCCTTAACCTTCAATGATTCACAAA 720
Qy 721 AGGAAGATGAAGTGATTCATGATTTAGAAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT 780
Db 721 AGGAAGATGAAGTGATTCATGATTTAGAAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT 780
Qy 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAACAAA 840
Db 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAACAAA 840
Qy 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAACATTTTGGATACATTTGATGTA 900
Db 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAATATCTAAACATTTTGGATACATTTGATGTA 900
Qy 901 ATGTAAGACTGTACAGACTTCTCTAGAAAACAGTTTGGGTTCATCTTTTCATTTCCCCAG 960
Db 901 ATGTAAGACTGTACAGACTTCTCTAGAAAACAGTTTGGGTTCATCTTTTCATTTCCCCAG 960
Qy 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAAATTGATT 1020
Db 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAAATTGATT 1020
Qy 1021 CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAATAATGAAGACCTTACTGATG 1080
Db 1021 CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAATAATGAAGACCTTACTGATG 1080
Qy 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACTACAGATAAATCGGGAACCTGTTAAACCAA 1140
Db 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACTACAGATAAATCGGGAACCTGTTAAACCAA 1140
Qy 1141 TTATGATGATGGCAACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAA 1200
Db 1141 TTATGATGATGGCAACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAA 1200
Qy 1201 ACAGTGTTCGGCTAAGTGATGCTCTTTTAAATAAAATTTGATTGGTGGTTACAGTCAAGCAA 1260
Db 1201 ACAGTGTTCGGCTAAGTGATGCTCTTTTAAATAAAATTTGATTGGTGGTTACAGTCAAGCAA 1260
Qy 1261 TTGAAGCGCTTCCCCCAGATAAAATATGGCCAAAATGAGAGTTTGTGCTAGAAATTCAGTGA 1320
Db 1261 TTGAAGCGCTTCCCCCAGATAAAATATGGCCAAAATGAGAGTTTGTGCTAGAAATTCAGTGA 1320
Qy 1321 GATTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGCTGACTACTTTCAATGG 1380
Db 1321 GATTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGCTGACTACTTTCAATGG 1380
Qy 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACGT 1440

Db 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACGT 1440
Qy 1441 CACAAGGTAAATCTCAAAAAAAGTAACAACACTTCTTCAAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAAATCTCAAAAAAAGTAACAACACTTCTTCAAAAAAGCTGTAGAACGTGGAGCAG 1500
Qy 1501 TACCACATAGAAATGCTGGAAATTTGCCCTGCGGAATTTTAAACCTCCAAAAAAGCAGCTGC 1560
Db 1501 TACCACATAGAAATGCTGGAAATTTGCCCTGCGGAATTTTAAACCTCCAAAAAAGCAGCTGC 1560
Qy 1561 TTTTCAGAGGAGGAAAAAAGAAATTTATCAGCATCTACCGTATTAACTGCCCAAGAAATCAT 1620
Db 1561 TTTTCAGAGGAGGAAAAAAGAAATTTATCAGCATCTACCGTATTAACTGCCCAAGAAATCAT 1620
Qy 1621 TTTCCGGTTCACTTGGGCATTTTACAGAAATAGGAACAACAGTTGTGATTCCAGAGGACAGA 1680
Db 1621 TTTCCGGTTCACTTGGGCATTTTACAGAAATAGGAACAACAGTTGTGATTCCAGAGGACAGA 1680
Qy 1681 CTACTAAAGCCAGGTTTATATATGGAGAAACATGCCACCAAGATGCAGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTATATATGGAGAAACATGCCACCAAGATGCAGAAATAGGTT 1740
Qy 1741 ACCGGAATTCATTGAGACAAACTAACAAAACTAAACAGTCAATGCCCATTTTGGAAAGATCC 1800
Db 1741 ACCGGAATTCATTGAGACAAACTAACAAAACTAAACAGTCAATGCCCATTTTGGAAAGATCC 1800
Qy 1801 CAGTTAAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Qy 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCCTGGATCTA 1920
Qy 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTTAAAGTCTGTTTCAAAATAGTCAAT 1980
Db 1921 AACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTTAAAGTCTGTTTCAAAATAGTCAAT 1980
Qy 1981 TCAAGGAACCTCTGGTGTGATGAAAGAGTTCTGAACTTTATTATTACTGATTCAATAA 2040
Db 1981 TCAAGGAACCTCTGGTGTGATGAAAGAGTTCTGAACTTTATTATTACTGATTCAATAA 2040
Qy 2041 CCTGAAAGAAATAAACCGGAATCAAGTCTTCTAGCTAAATTTAGAGAAACTAAAGAGTATC 2100
Db 2041 CCTGAAAGAAATAAACCGGAATCAAGTCTTCTAGCTAAATTTAGAGAAACTAAAGAGTATC 2100
Qy 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACAGAGTAAAGTGGCAAGTAAAGAGTAAAGTCAAGT 2160
Db 2101 AAGAACCCAGAGGTTCCAGAGAGTAAACAGAGTAAAGTGGCAAGTAAAGAGTAAAGTCAAGT 2160
Qy 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAA 2220
Qy 2221 AAGTTAATACAGAGCAGAAACATACCCTTTTGGAGCAACCTGTCTTTTCAAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCCTTTTGGAGCAACCTGTCTTTTCAAAAAC 2280
Qy 2281 AGTCACCAACCAATATCAACATCTAAATGGTTGACCCCAAAATCTATTGTAAAGACACCAA 2340
Db 2281 AGTCACCAACCAATATCAACATCTAAATGGTTGACCCCAAAATCTATTGTAAAGACACCAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAGAACTCCAGTTGTTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAGAACTCCAGTTGTTAAAGAAATGACT 2400
Qy 2401 TTCCACCTGCTTGTGCTGTCAGTGTCAACACCTTATGGCCAAACCTGCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGCTGTCAGTGTCAACACCTTATGGCCAAACCTGCTGTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAAATACTTGGCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATACTTGGCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAA 2520

QY 2521 ATGAATGCATTTTCGGTTAAAGGAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
Db |||||
QY 2521 ATGAATGCATTTTCGGTTAAAGGAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
Db |||||
QY 2581 GTTCAAGCAGAGGTATTTTCAGGTTTAAATGAAAGAAACAGATATATGCTATAAAATATG 2640
Db |||||
QY 2581 GTTCAAGCAGAGGTATTTTCAGGTTTAAATGAAAGAAACAGATATATGCTATAAAATATG 2640
Db |||||
QY 2641 TGAACCTTAGAAGACAGATACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTTATT 2700
Db |||||
QY 2641 TGAACCTTAGAAGACAGATACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTTATT 2700
Db |||||
QY 2701 TGAATAAATACTAACACACACAGTGAAGATCATCCGACTTTATGATTAATGAAATCAGG 2760
Db |||||
QY 2701 TGAATAAATACTAACACACACAGTGAAGATCATCCGACTTTATGATTAATGAAATCAGG 2760
Db |||||
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAAA 2820
Db |||||
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAAA 2820
Db |||||
QY 2821 AGAAAAATCCATTGATCCATGGGAACGCAAGATTACTGGAATAATGTTAGAGGCAG 2880
Db |||||
QY 2821 AGAAAAATCCATTGATCCATGGGAACGCAAGATTACTGGAATAATGTTAGAGGCAG 2880
Db |||||
QY 2881 TTCACACAATCCATCAACATGGCATTTGTCACAGTGTCTTAAACCAAGTAACTTTCTGA 2940
Db |||||
QY 2881 TTCACACAATCCATCAACATGGCATTTGTCACAGTGTCTTAAACCAAGTAACTTTCTGA 2940
Db |||||
QY 2941 TAGTTGATGGAATGCTAAAGTAAATGATTTTGGGATGCAACCAAAATGCAACCCAGATA 3000
Db |||||
QY 2941 TAGTTGATGGAATGCTAAAGTAAATGATTTTGGGATGCAACCAAAATGCAACCCAGATA 3000
Db |||||
QY 3001 CAACAAGTGTGTTTAAAGATTCTCAGGTTGGCAGATTAAATATATGCCACCAAGCAAA 3060
Db |||||
QY 3001 CAACAAGTGTGTTTAAAGATTCTCAGGTTGGCAGATTAAATATATGCCACCAAGCAAA 3060
Db |||||
QY 3061 TCAAGATATGTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATAAGCCCCAAA 3120
Db |||||
QY 3061 TCAAGATATGTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAGATAAGCCCCAAA 3120
Db |||||
QY 3121 GTGATGTTTGGTCTCTTAGGATGTTATTTTGTACTATATGACTTACCGGAAACACCATTTT 3180
Db |||||
QY 3121 GTGATGTTTGGTCTCTTAGGATGTTATTTTGTACTATATGACTTACCGGAAACACCATTTT 3180
Db |||||
QY 3181 AGCAGATAATTAATCAGATTCTTAAATTACATGCCATAATTGATCCTTAATCATGAATTG 3240
Db |||||
QY 3181 AGCAGATAATTAATCAGATTCTTAAATTACATGCCATAATTGATCCTTAATCATGAATTG 3240
Db |||||
QY 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTGTTAAAGTGTGTTTAAAGGG 3300
Db |||||
QY 3241 AATTTCCGATATTTCCAGAGAAAGATCTTCAAGATGTTGTTAAAGTGTGTTTAAAGGG 3300
Db |||||
QY 3301 ACCCAAAACAGAGGATATCCATTCTCGCTCCTGAGTGTGTTAAAGTGTGTTTAAAGGG 3360
Db |||||
QY 3301 ACCCAAAACAGAGGATATCCATTCTCGCTCCTGAGTGTGTTAAAGTGTGTTTAAAGGG 3360
Db |||||
QY 3361 CTCATCCAGTTAACCAATGGCCCAAGGACCACTGAGGAAATGAAATATGTTCTGGGCC 3420
Db |||||
QY 3361 CTCATCCAGTTAACCAATGGCCCAAGGACCACTGAGGAAATGAAATATGTTCTGGGCC 3420
Db |||||
QY 3421 AACTTGTGTTGCTGAATCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3480
Db |||||
QY 3421 AACTTGTGTTGCTGAATCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3480
Db |||||
QY 3481 ACTATAGTGGTGGTGAAGTCAATATCTTCTCATCTCCAGACTTTTGAATAAAAGGG 3540
Db |||||
QY 3481 ACTATAGTGGTGGTGAAGTCAATATCTTCTCATCTCCAGACTTTTGAATAAAAGGG 3540
Db |||||
QY 3541 GAAAAAATGATTTGAGTTATTCGTAATGTCAGATAGGAGTATATAATATGACT 3600
Db |||||
QY 3541 GAAAAAATGATTTGAGTTATTCGTAATGTCAGATAGGAGTATATAATATGACT 3600
Db |||||

QY 3601 GTTATACCTTGAATCCCTGTGAAATCTACATTTGAAGACAAATCACTCTGAAGTGT 3660
Db |||||
QY 3601 GTTATACCTTGAATCCCTGTGAAATCTACATTTGAAGACAAATCACTCTGAAGTGT 3660
Db |||||
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAAACTGTAAAAATAGCAACCAC 3720
Db |||||
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAAACTGTAAAAATAGCAACCAC 3720
Db |||||
QY 3721 TTATGGCACTGTATATATTTAGACTTGTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
Db |||||
QY 3721 TTATGGCACTGTATATATTTAGACTTGTCTCTGTTTATGCTCTTGTGTAATCTAC 3780
Db |||||
QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTTCTAAAAAATCTTG 3840
Db |||||
QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGTGGATAGCAAGTATATTTCTAAAAAATCTTG 3840
Db |||||
QY 3841 TAAATAAAGTTTTGTGGCTAAAAATGA 3866
Db |||||
QY 3841 TAAATAAAGTTTTGTGGCTAAAAATGA 3866
Db |||||

RESULT 12

US-10-295-027-481
; Sequence 481, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 481
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-481

Query Match 66.5%; Score 2569.2; DB 16; Length 2574;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 978 ATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAAATTGATTTCCATAATGAACAAAGTG 1037
Db |||||
1 ATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAAATTGATTTCCATAATGAACAAAGTG 60
QY 1038 AGAGACATTTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAA 1097
Db |||||
61 AGAGACATTTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAA 120
QY 1098 ATTTCTGCTGATACTACAGATAAAGTGGGAACTGTTAAACCAAAATATGATGATGGCAAAAC 1157
Db |||||
121 ATTTCTGCTGATACTACAGATAAAGTGGGAACTGTTAAACCAAAATATGATGATGGCAAAAC 180
QY 1158 AACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTCCGCTAAGT 1217
Db |||||
181 AACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTCCGCTAAGT 240
QY 1218 GATGCTCTTTTAAATAAAATGATTTGGTTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCCCA 1277
Db |||||
241 GATGCTCTTTTAAATAAAATGATTTGGTTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCCCA 300
QY 1278 GATAAATATGGCCAAAAATGAGAGTTTGTGCTAGAAATTTCAAGTGAGATTTGCTGAATTAATA 1337
Db |||||
301 GATAAATATGGCCAAAAATGAGAGTTTGTGCTAGAAATTTCAAGTGAGATTTGCTGAATTAATA 360
QY 1338 GCTATTCAAGAGCCAGATGATGCACGTGACTACTTTTCAAAATGCCAGAGCAAACTGCAAG 1397
Db |||||
361 GCTATTCAAGAGCCAGATGATGCACGTGACTACTTTTCAAAATGCCAGAGCAAACTGCAAG 420
QY 1398 AAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACCTGCACAAAGGTAATGCAAA 1457
Db |||||
421 AAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACCTGCACAAAGGTAATGCAAA 480
QY 1458 AAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCACCTAGAAATGCTG 1517
Db |||||
481 AAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCACCTAGAAATGCTG 540
QY 1518 GAAATTTGCCCTGCGGAATTTAAACCTCAAAAAAAGCAGCTGCTTTCAGAGGAGGAAAAAG 1577
Db |||||
541 GAAATTTGCCCTGCGGAATTTAAACCTCAAAAAAAGCAGCTGCTTTCAGAGGAGGAAAAAG 600
QY 1578 AAGAAATTTATCAGCATCTACGGTATTTAACTGCCCCAAGAATCATTTTCCGGTTCACTTGGG 1637
Db |||||
601 AAGAAATTTATCAGCATCTACGGTATTTAACTGCCCCAAGAATCATTTTCCGGTTCACTTGGG 660
QY 1638 CATTTACAGAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTT 1697
Db |||||
661 CATTTACAGAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTT 720
QY 1698 TTATATGGAGAGAACATGCCACCACCAAGATGCAGAAATAGGTTACCGGAATTCATTTGAGA 1757
Db |||||
721 TTATATGGAGAGAACATGCCACCACCAAGATGCAGAAATAGGTTACCGGAATTCATTTGAGA 780
QY 1758 CAAACTAAACAACTTAAACAGTCAATGCCATTTGGAAGAGTCCCAGTTAACCTTCTAAAT 1817
Db |||||
781 CAAACTAAACAACTTAAACAGTCAATGCCATTTGGAAGAGTCCCAGTTAACCTTCTAAAT 840
QY 1818 AGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGAAAGACAA 1877
Db |||||
841 AGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGAAAGACAA 900
QY 1878 ACCTCTAGATCAGAAATCCGAGATTTGGTTGTGCTTGAATCTTAAACCAAGTGGAAATGAT 1937
Db |||||
901 ACCTCTAGATCAGAAATCCGAGATTTGGTTGTGCTTGAATCTTAAACCAAGTGGAAATGAT 960
QY 1938 TCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACCTCTGGTG 1997
Db |||||
961 TCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACCTCTGGTG 1020
QY 1998 TCAGATGAAAAGAGTTCTGAACTTATTTACTGATTCATTAACCTTGAAGAAATAAAACG 2057
Db |||||
1021 TCAGATGAAAAGAGTTCTGAACTTATTTACTGATTCATTAACCTTGAAGAAATAAAACG 1080

QY 2058 GAATCAAGTCTTCTAGCTAAATTAGAAGAAACTAAAGAGTATCAAGAAACCAGAGGTTCCA 2117
Db |||||
1081 GAATCAAGTCTTCTAGCTAAATTAGAAGAAACTAAAGAGTATCAAGAAACCAGAGGTTCCA 1140
QY 2118 GAGAGTAACCCAGAAACAGTGGCAAGCTAAGAGAAAAAGTCAGAGTGTATTAACCCAGAAATCCT 2177
Db |||||
1141 GAGAGTAACCCAGAAACAGTGGCAATCTAAGAGAAAAAGTCAGAGTGTATTAACCCAGAAATCCT 1200
QY 2178 GCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAAAAAGTTAATACAGAGCAG 2237
Db |||||
1201 GCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAAAAAGTTAATACAGAGCAG 1260
QY 2238 AAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAACAGTCACCACCAATATCA 2297
Db |||||
1261 AAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAACAGTCACCACCAATATCA 1320
QY 2298 ACATCTAAATGGTTTGACCCCAAAATCTATTTGTAAAGACACCAAGCAGCAATACCTTGGAT 2357
Db |||||
1321 ACATCTAAATGGTTTGACCCCAAAATCTATTTGTAAAGACACCAAGCAGCAATACCTTGGAT 1380
QY 2358 GATTACATGAGCTGTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTCTAG 2417
Db |||||
1381 GATTACATGAGCTGTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTCTAG 1440
QY 2418 TTGTCAACACCTTTATGGCCAACTGCTGCTTTCCAGCAGCAACAGCATCAAAATACTTTGCC 2477
Db |||||
1441 TTGTCAACACCTTTATGGCCAACTGCTGCTTTCCAGCAGCAACAGCATCAAAATACTTTGCC 1500
QY 2478 ACTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAATGAATGCATTTCCGTT 2537
Db |||||
1501 ACTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAATGAATGCATTTCCGTT 1560
QY 2538 AAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAAGTGGAGTTCAAGCAAGGTATTT 2597
Db |||||
1561 AAAGGAAGAAATTTATTCATATTTAAAGCAGATAGGAAAGTGGAGTTCAAGCAAGGTATTT 1620
QY 2598 CAGGTGTTAAATGAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGAAGCA 2657
Db |||||
1621 CAGGTGTTAAATGAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGAAGCA 1680
QY 2658 GATAACCAAACTCTTGATAGTTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACAA 2717
Db |||||
1681 GATAACCAAACTCTTGATAGTTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACAA 1740
QY 2718 CACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCACGGACCAAGTACATCTACATG 2777
Db |||||
1741 CACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCACGGACCAAGTACATCTACATG 1800
QY 2778 GTPAATGGAGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAAGAAAAAATCCATTTGAT 2837
Db |||||
1801 GTPAATGGAGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAAGAAAAAATCCATTTGAT 1860
QY 2838 CCATGGGAACCGCAAGAGTTACTGGAAAAATATGTTAGAGGCAGTTTACACAAATCCATCAA 2897
Db |||||
1861 CCATGGGAACCGCAAGAGTTACTGGAAAAATATGTTAGAGGCAGTTTACACAAATCCATCAA 1920
QY 2898 CATGGCATTTGTTCACTAGTATCTTAAACCAAGTAACTTTCTGATAGTTGGAATGCTA 2957
Db |||||
1921 CATGGCATTTGTTCACTAGTATCTTAAACCAAGTAACTTTCTGATAGTTGGAATGCTA 1980
QY 2958 AAGCTAATTTGATTTTGGGATTTGCAACCAAAATGCAACCAAGTAACTTTGTTTAA 3017
Db |||||
1981 AAGCTAATTTGATTTTGGGATTTGCAACCAAAATGCAACCAAGTAACTTTGTTTAA 2040
QY 3018 GATTCTCAGGTTGGCAGAGTTAATTTATATGCGCACCAAGCAATCAAAAGATATGTCTTCC 3077
Db |||||
2041 GATTCTCAGGTTGGCAGAGTTAATTTATATGCGCACCAAGCAATCAAAAGATATGTCTTCC 2100
QY 3078 TCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAAGTATGTTTGGTCCCTTA 3137
Db |||||
2101 TCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAAGTATGTTTGGTCCCTTA 2160
QY 3138 GGATGATTTTGTACTATATGACTTACGGGAAAAACCACTTTTCAGCAGATAATTAATCAG 3197

Db 2161 GGATGTAATTTGTACTATATGACTTACGGGAAAACACCATTTTCAGCAGATAATTAATCAG 2220
QY 3198 ATTTCTAAATACATGCCATAATGATCCTAATCATGAAATGAATTTCCCGATATTTCCA 3257
Db 2221 ATTTCTAAATACATGCCATAATGATCCTAATCATGAAATGAATTTCCCGATATTTCCA 2280
QY 3258 GAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGACCCAAAACAGAGGATA 3317
Db 2281 GAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGAGGACCCAAAACAGAGGATA 2340
QY 3318 TCCATTTCTGAGCTCCTGGCTCATCCATATGTTCAAATTTCAAATCATCCAGTTAAACAA 3377
Db 2341 TCCATTTCTGAGCTCCTGGCTCATCCCTATGTTCAAATTTCAAATCATCCAGTTAAACAA 2400
QY 3378 ATGGCCAAAGGAACCACTGAAGAAATGAAATATGTTCTGGGCCAACTTGGTCTGAAT 3437
Db 2401 ATGGCCAAAGGAACCACTGAAGAAATGAAATATGTTCTGGGCCAACTTGGTCTGAAT 2460
QY 3438 TCTCCTAACTCCATTTTGAAGCTGTAAACCTTTATATGAACACTATAGTGGTGGA 3497
Db 2461 TCTCCTAACTCCATTTTGAAGCTGTAAACCTTTATATGAACACTATAGTGGTGGA 2520
QY 3498 AGTCATAATTTCTCATCCTCCAAGACTTTTGAAGAAAGGAGGAAATGA 3551
Db 2521 AGTCATAATTTCTCATCCTCCAAGACTTTTGAAGAAAGGAGGAAATGA 2574

RESULT 13

US-10-173-999-33
; Sequence 33, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-173-999-33

Query Match 66.5%; Score 2569.2; DB 16; Length 2574;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 978 ATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAAATTGATTCCTAATAATGAACAAAGTG 1037
Db 1 ATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAAATTGATTCCTAATAATGAACAAAGTG 60
QY 1038 AGAGACATTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAATAAGCTTGAATAAA 1097
Db 61 AGAGACATTAAAAATAAGTTTAAAAATGAAGACCTTACTGATGAATAAGCTTGAATAAA 120
QY 1098 ATTTCTGCTGATCTACAGATAAATCGGGAACTGTTAAACCAATTTATGATGATGGCAAC 1157
Db 121 ATTTCTGCTGATCTACAGATAAATCGGGAACTGTTAAACCAATTTATGATGATGGCAAC 180

QY 1158 AACCCAGAGGACTGGTTGAGTTTGTGTTGCTCAAACTAGAGAAAAACAGTGTTCGCTAAGT 1217
Db 181 AACCCAGAGGACTGGTTGAGTTTGTGTTGCTCAAACTAGAGAAAAACAGTGTTCGCTAAGT 240
QY 1218 GATGCTCTTTTAAATAAATGATTTGGTTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCCCA 1277
Db 241 GATGCTCTTTTAAATAAATGATTTGGTTCGTTACAGTCAAGCAATTTGAAGCGCTTCCCCCA 300
QY 1278 GATAAATATGGCCCAAAATGAGAGTTTTTGGTCTAGAAATTTCAAGTGAGATTTGCTGAATTAATA 1337
Db 301 GATAAATATGGCCCAAAATGAGAGTTTTTGGTCTAGAAATTTCAAGTGAGATTTGCTGAATTAATA 360
QY 1338 GCTATTCAAGAGCCAGATGATGCACGTGACTCTTTCAAATGGCCAGAGCAAACTGCAAG 1397
Db 361 GCTATTCAAGAGCCAGATGATGCACGTGACTCTTTCAAATGGCCAGAGCAAACTGCAAG 420
QY 1398 AAATTTGCTTTTGTTCATATATCTTTTGGCACAATTTGAACTGTCAAGGTAAATGTCAA 1457
Db 421 AAATTTGCTTTTGTTCATATATCTTTTGGCACAATTTGAACTGTCAAGGTAAATGTCAA 480
QY 1458 AAAAGTAAACAACTTCTTCAAAAAAGCTGTAGAACGTGGAGCAGTACCACCTAGAAATGCTG 1517
Db 481 AAAAGTAAACAACTTCTTCAAAAAAGCTGTAGAACGTGGAGCAGTACCACCTAGAAATGCTG 540
QY 1518 GAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTCAAGAGGAGGAAAAG 1577
Db 541 GAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTCAAGAGGAGGAAAAG 600
QY 1578 AAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAATTTCCCGGTTCACTTGGG 1637
Db 601 AAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAATTTCCCGGTTCACTTGGG 660
QY 1638 CATTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT 1697
Db 661 CATTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT 720
QY 1698 TTATATGGAGAGAACATGCCACCACAAGATGCAAGATGAGAAATAGGTTACCGGAATTCATTGAGA 1757
Db 721 TTATATGGAGAGAACATGCCACCACAAGATGCAAGATGAGAAATAGGTTACCGGAATTCATTGAGA 780
QY 1758 CAACTAAACAAACTAAACAGTCAATGCCCAATTTGGAAGAGTCCCAGTTAACCTTCTAAAT 1817
Db 781 CAACTAAACAAACTAAACAGTCAATGCCCAATTTGGAAGAGTCCCAGTTAACCTTCTAAAT 840
QY 1818 AGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGAAAAGACAA 1877
Db 841 AGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGAAAAGACAA 900
QY 1878 ACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCTGGATCTAAACCAAGTGGAATGAT 1937
Db 901 ACCTCTAGATCAGAAATGCCGAGATTTGGTTGTCCTGGATCTAAACCAAGTGGAATGAT 960
QY 1938 TCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTCAGGAACTCTCTGGTG 1997
Db 961 TCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTCAGGAACTCTCTGGTG 1020
QY 1998 TCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCAATAACCTGAAAGATAAAACG 2057
Db 1021 TCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCAATAACCTGAAAGATAAAACG 1080
QY 2058 GAATCAAGTCTTCTAGCTAAATTAAGAGAACTAAAGAGTATCAAGAACCCAGAGGTTCCA 2117
Db 1081 GAATCAAGTCTTCTAGCTAAATTAAGAGAACTAAAGAGTATCAAGAACCCAGAGGTTCCA 1140
QY 2118 GAGAGTAAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAAGAGTGTATTAAACAGAAATCCT 2177
Db 1141 GAGAGTAAACAGAAAACAGTGGCAATCTAAGAGAAAGTCAAGAGTGTATTAAACAGAAATCCT 1200
QY 2178 GCTGCATCTTCAAATCACTGGCAGATTCGGAGTTAGCCCGAGTTAGCCCGAAAGTTAATACAGAGCAG 2237
Db 1201 GCTGCATCTTCAAATCACTGGCAGATTCGGAGTTAGCCCGAGTTAGCCCGAAAGTTAATACAGAGCAG 1260

QY 2238 AAACATACCACCTTTTGGACCAACCTGTCTTTTCAGTTTCAAAAACAGTCACCACCAATATCA 2297
Db 1261 AAACATACCACCTTTTGGACCAACCTGTCTTTTCAGTTTCAAAAACAGTCACCACCAATATCA 1320
QY 2298 ACATCTAAATGGTTGACCCCAAAATCTATTTGTAAGACACCAAGCAGCAATACCTTGGAT 2357
Db 1321 ACATCTAAATGGTTGACCCCAAAATCTATTTGTAAGACACCAAGCAGCAATACCTTGGAT 1380
QY 2358 GATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTGAG 2417
Db 1381 GATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTGAG 1440
QY 2418 TTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAACAGCATCAAATACCTGCC 2477
Db 1441 TTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAACAGCATCAAATACCTGCC 1500
QY 2478 ACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAAAATGAATGCATTTTCGGTT 2537
Db 1501 ACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAAAATGAATGCATTTTCGGTT 1560
QY 2538 AAAGGAAGAAATTTATCCATATTTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGGTATTT 2597
Db 1561 AAAGGAAGAAATTTATCCATTTTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGGTATTT 1620
QY 2598 CAGGTGTTAAATGAAAAAGAAACAGATATATGCTATATAAAATATGTAACTTAGAAGAGCA 2657
Db 1621 CAGGTGTTAAATGAAAAAGAAACAGATATATGCTATATAAAATATGTAACTTAGAAGAGCA 1680
QY 2658 GATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACAA 2717
Db 1681 GATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACAA 1740
QY 2718 CACAGTGATAAGATCATCCGACTTTATGATATGAAATCAGGACCAAGTACATCTACATG 2777
Db 1741 CACAGTGATAAGATCATCCGACTTTATGATATGAAATCAGGACCAAGTACATCTACATG 1800
QY 2778 GTAATGGAGTGTGGAATATTGATCTTAAATAGTTGGCTTAAAGAGAAAAATCCATTGAT 2837
Db 1801 GTAATGGAGTGTGGAATATTGATCTTAAATAGTTGGCTTAAAGAGAAAAATCCATTGAT 1860
QY 2838 CCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAAGTTCACACAATCCATCAA 2897
Db 1861 CCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAAGTTCACACAATCCATCAA 1920
QY 2898 CATGGCATTTGTTACAGTGATCTTAAACCAAGTAACTTCTGATAGTTGATGGAATGCTA 2957
Db 1921 CATGGCATTTGTTACAGTGATCTTAAACCAAGTAACTTCTGATAGTTGATGGAATGCTA 1980
QY 2958 AAGCTAAATGATTTTGGGATTCGAAACCAATGCAACCAAGATACAAAGTGTGTTTAA 3017
Db 1981 AAGCTAAATGATTTTGGGATTCGAAACCAATGCAACCAAGATACAAAGTGTGTTTAA 2040
QY 3018 GATTTCTCAGGTTGGCAGATTAATTAATATGCAACCAAGCAATCAAAAGATATGCTTCC 3077
Db 2041 GATTTCTCAGGTTGGCAGATTAATTAATATGCAACCAAGCAATCAAAAGATATGCTTCC 2100
QY 3078 TCCAGAGAGAATGGGAAATCTAAGTCAAAAGATAAGCCCCCAAAAGTATGTTGGTCCCTTA 3137
Db 2101 TCCAGAGAGAATGGGAAATCTAAGTCAAAAGATAAGCCCCCAAAAGTATGTTGGTCCCTTA 2160
QY 3138 GGATGTATTTTGTACTATATGACTTACGGGAAAAACACCATTTTCAGCAGATAATTAATCAG 3197
Db 2161 GGATGTATTTTGTACTATATGACTTACGGGAAAAACACCATTTTCAGCAGATAATTAATCAG 2220
QY 3198 ATTTCTAAATACATGCCATAATTTGATCTCTAATCATGAAATTTCCCGATATTTCCA 3257
Db 2221 ATTTCTAAATACATGCCATAATTTGATCTCTAATCATGAAATTTCCCGATATTTCCA 2280
QY 3258 GAGAAAGATCTTCAAGATGTGTTTAAAGTGTGTTTAAAGAGGACCCCAAAACAGAGGATA 3317
Db 2281 GAGAAAGATCTTCAAGATGTGTTTAAAGTGTGTTTAAAGAGGACCCCAAAACAGAGGATA 2340
QY 3318 TCCATTCCTGAGCTCCTGGCTCATCCATATGTTCAAAATTTCAAACCTCATCCAGTTAACCAA 3377

Db 2341 TCCATTCCTGAGCTCCTGGCTCATCCCTATGTTCAAAATTCAAACTCATCCAGTTAACCAA 2400
QY 3378 ATGGCCAAGGGAACCACTGAAGAAATGAATATGTTCTGGCCAACTTGTGTTGTTCTGAAT 3437
Db 2401 ATGGCCAAGGGAACCACTGAAGAAATGAATATGTTCTGGCCAACTTGTGTTGTTCTGAAT 2460
QY 3438 TCTCCTAACTCCATTTTGAAGCTGCTAAACTTTATATGAACACTATAGTGGTGGA 3497
Db 2461 TCTCCTAACTCCATTTTGAAGCTGCTAAACTTTATATGAACACTATAGTGGTGGA 2520
QY 3498 AGTCATAAATTTCTCATCTCCAAAGACTTTTGAAGAAAAAAGGGGAAAAAATGA 3551
Db 2521 AGTCATAAATTTCTCATCTCCAAAGACTTTTGAAGAAAAAAGGGGAAAAAATGA 2574

RESULT 14

US-10-027-632-110237
; Sequence 110237, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 110237
; LENGTH: 2470
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-110237

Query Match 10.4%; Score 402; DB 13; Length 2470;
Best Local Similarity 79.0%; Pred. No. 6.6e-86;
Matches 518; Conservative 0; Mismatches 130; Indels 8; Gaps 3;

QY 9 TTTTCTTTTCTTTTGTGAGATGGAGTTTCACTCTGTTGGCCAGGCTGGAGTGAATGGCA 68
Db 1178 TTTCTTTTCTTTTGTGAGCGGAGTTTCACTCTGTTGGCCAGGCTGGAGTGAATGGCA 1237
QY 69 CAATCTCAGCTTACTGCAACCTCCGCCCTCCCGGTTCAAGCGATTCTCCTGCCCTCAGCCT 128
Db 1238 CGAATTCGGCTCACTGCAACCTCCACCTCCCGGTTCAAGCGATTCTCCTGCCCTCAGCCT 1297
QY 129 CTCAAGTAGCTGGGATTACAGGCATGTGCCACCAACCCCTGGCTAACTAATTTCTTTTCTA 188
Db 1298 CCCAAGTAGCTGGAATTACAGGCACGTGTACCACTCCTGGCTAA-----TTTGTAGT 1351
QY 189 TTTAGTAGAGATGGGGTTTCAACCATGTTGGTCAGGCTGGTCTTGAACCTCCTGACCTCAGG 248
Db 1352 TCTAGTAGAGACAGGGTTTCCCAACGTTGGTCAGGCTAGTTTCGAACACCCGACCTCAGG 1411
QY 249 TGATCCACTTGCTTGGCCCTCCCAAGTGTAGGATTACAGCCGTGAAGTGTGCCTGGC 308
Db 1412 TGATCCACCCACCTCAGCCTTCCCAAGTGTAGGATTACAGGCATGAGCCACTGCACCTG 1471

Db 642 GGACAAAGAGAGTGTGTGAGCATCGACAGTACTAAGTGCCCAAGAGCCGTTTCTCCAGCTC 701

Qy 1631 ACTTGGGCATTTACAGAAATAGGAACAACAGATTGTGATTCAGAGGACAGACTACTAAAGC 1690

Db 702 ACTTGGAAATGTACAGAAATCGGAGCATCAGCTGTGAGTCCAGAGGACAGGCTGGGCAGC 761

Qy 1691 CAGGTTTTTATATGAGAGAGAACATGCCACCACAAAGATGCAGAAATAGGTTACCGGAATTC 1750

Db 762 CAGGTTTTTATATGAGAGAACTTGCCTCCACAAGATGCCGAAGTGAGGCATCAAAACCC 821

Qy 1751 ATTGAGACAAACTAAACAAACTAAACAGTCATGCCCATTTTGGAAAGATGCCAGTTAACTT 1810

Db 822 CTTCAAGCAGACTCAGCGAGCTAAACGTCATGCCCTTTTGAAGAGTCCCAGTCAATCT 881

Qy 1811 TCTAAATAGCCCGAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATTATGAA 1870

Db 882 TCTAAACAGCCCGAGATTCTATGTGAAGACAGATAGCTCAGCTGTGACACAGTTAAACAAC 941

Qy 1871 AAGACAAACCTCTAGATCAGAAATGCCCGAGATTGGTTGTGCTGGATCTAAACCAAGTGG 1930

Db 942 AAGAAAGGGTCAAGACCAGA--CCGAGACGGGATTCTGCCCGGCTCCAGACCAGCTGG 998

Qy 1931 AAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAATAGTCATTTCAAGGAACC 1990

Db 999 CAGTGATTCCCTATGAACCTGAGAGGTTTAAAGCCCATTCAAACTATCTATTTGAAGACTC 1058

Qy 1991 TCTGGTGTGAGTGAAGAGATTCTGAACCTTATTATTAAGTCAATTAACCCCTGAAGAA 2050

Db 1059 TTTGGTGTCCAATGAAAGAGTTCTGAAC--TTATGTCTGATTTAATAGCCTTGAAGAG 1115

Qy 2051 TAAACCGGAATCAAGTCTTCTAGCTAAATTAAGAAAGAACTAAAGAGTATCAAGAACCCAGA 2110

Db 1116 TAAACAGATTCAAG--TCTAACAAATTTGAAGAACTA-----AGCCAGA 1160

Qy 2111 GGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAAGAGAAAGTCAGAGTGATTAACCA 2170

Db 1161 GATTGCAGAAAGAGGCCCATGTCAGTGGCAGTCTACCAGAAAGCCCGAGTGTGTCTCA 1220

Qy 2171 GAATCCTGCTGCATCTTCAAATCACTGSCAGATTCCGGAGTTAGCCCGAAAGTTAATAC 2230

Db 1221 GAACCTGCTGCTTTCACCCCTGCGGCAGCTTCCAGATGTCACCCCGAAGSCTGA--- 1277

Qy 2231 AGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAACAGTCACCACC 2290

Db 1278 -----CAAAGATCACACC 1292

Qy 2291 AATATCAACATCTAAATGGTTTGACCCCAAAATCTATTTGTAAAGACACCAAGCAGCAATAC 2350

Db 1293 AATATCAGTTCCTAAATGGCTTGATCCCAAGTCTGCTTGTGACACACCTAGTAGCAGCTC 1352

Qy 2351 CTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGC 2410

Db 1353 CTTGGATGATTACATGAAATGTTTTAAGACTCCAGTTGTAAAGAAATGACTTTCCACCTGC 1412

Qy 2411 TTGTCAGTTGTCAACACCTTATGGCCAACTGCTGCTTTCCAGCAGCAACAGCATCAAAAT 2470

Db 1413 CTGTCCATCATCAACACCTTACAGCCAGCTTGCCCGCTCCAGCAGCAACAGCAGAGGG 1472

Qy 2471 ACTTGCCACTCCACTTCAAAATTTACAGGTTTGTAGCATCTTCTTCAGCAAAATGAATGCAT 2530

Db 1473 ACTCAGCACTCCTCTTCAAGCTTGCAGATTTTCAGGTTCTTCATCAATAAATGAATGCAT 1532

Qy 2531 TTCGGTTAAAGGAAGAAATTTATCCATATTAAGCAGATAGGAAGTGGAGTTTCAAGCAA 2590

Db 1533 TTCAGTTAACGGAAGAAATTTATCCATATTAAGCAGATAGGAGTGGAGTTTCCAGTAA 1592

Qy 2591 GGTATTTCAAGGTGTTAAATGAAAGAAACAGATATATGCTATAAAATATGTGAACCTAGA 2650

Db 1593 GGTGTTTCAGGTATTGAATGAGAAAGAAACAGATAAAACGCTATCAAAATATGTGAACCTAGA 1652

Qy 2651 AGAAGCAGATAACCAAACTCTTGTAGTTACCGGAACGAAATAGCTTATTTGAATAAACT 2710

Db 1653 AGACGCCGATAGCCCAAACTATTGAGAGCTACCGCAACGAGATACGGTTTTTTGAACAACT 1712

Qy 2711 ACAACAAACACAGTGTAAAGATCATCCGACTTTATGATTATGAATCAGGACCAGTACAT 2770

Db 1713 ACAGCAACACAGTGTAAAGATCATCCGCTTCTATGATTATGAATCAGGACGATACAT 1772

Qy 2771 CTACATGGTAATGGAGTGTGGAATATTTGATCTTAATAGTTGGCTTAAAAAGAAAAATC 2830

Db 1773 CTACATGGTAATGGAAATGTGGAACATTGACCTAAATAGTTGGCTTAAAAAGAAAAATC 1832

Qy 2831 CATTGATCCATGGGAACGCAAGAGTTTACTGGAATAATATGTTAGAGGCAGTTTCACACAAT 2890

Db 1833 CATCAATCCATGGGAACGCAAGAGCTACTGGAATAAATCATGTTGGAGCAGTACACATAAT 1892

Qy 2891 CCATCAACATGGCAATTTGTTTCAAGTGTATTAACCCAGTAACTTTCTGATAGTTGATGG 2950

Db 1893 CCATCAGCATGGTATTGTTTCATAGTGAICTGAAGCCTGCTAACTTTGTGATAGTGGATGG 1952

Qy 2951 AATGCTAAAGCTAATTCATTTTGGGATTGCAAAACCAATGCAACAGATACAACAAGTGT 3010

Db 1953 AATGCTAAAGCTAATTCATTTTGGGATTGCAAAACCAATGCAACAGATACAACAAGCAT 2012

Qy 3011 TGTTAAAGATTCTCAGTTGGCACAGTTAATATATATGCCACCAGAAAGCAATCAAGATAT 3070

Db 2013 TGTTAAAGATTCTCAGTTGGCACAGTTAATATATATGCCCCCAGAAAGCAATCAGAGACAT 2072

Qy 3071 GTCTTCTCCAGAGAGAAATCTAAGTCAAAAGATAAGCCCCAAAAGTGATGTTTG 3130

Db 2073 GTCTTCTTCAAGAGAAATTCGAAATCAGGACCAAGGTAAGTCCAGAGAGTGATGCTG 2132

Qy 3131 GTCTTAGGATGATTTTGTACTATATGACTTACGGGAAACACCATTTTTCAGCAGATAAT 3190

Db 2133 GTCTTGGGTGCATTTTGTACTACATGACTTATGGGAGGAGCCCATTTTCAGCAGATCAT 2192

Qy 3191 TAATCAGATTTCTAATTACATGCCATAATTTGATCCTTAATCATGAATTTGAATTTCCCGA 3250

Db 2193 CAATCAGGTCTCTAAACTGCACGCCATATCAACCTGTCTCATGAGATTGAATTTCCCGA 2252

Qy 3251 TATTCAGAGAAAGATCTTCAAGATGTGTAAAGTGTGTTTAAAGGGACCCAAACA 3310

Db 2253 GATTCGGAAAAAGATCTTCGAGACGTGTAAAGTGTGTTTAGTGAAGAACCCCTAAAGA 2312

Qy 3311 GAGGATATCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAACTCATCCAGT 3370

Db 2313 GAGGATATCTATCCCTGAGCTTCTCACACATCCGTATGTTCAAATTCAGCCCCATCCAGG 2372

Qy 3371 TAACCAAAATGGCCAAAGGAACCACTGAAGAAATGAATAATATGTTCTGGGCCAACTTGTGG 3430

Db 2373 CAGCCAAATGGCTAGGGAGCCACTGATGAATAAATAATATGTTGGGTCAACTTGTGG 2432

Qy 3431 TCTGAATTCCTTAACCTCCATTTTGAAGAGTGTGCTAAA 3467

Db 2433 TCTGAATTCCTTAACCTCCATCTTGAAGAACTGCAAAA 2469

RESULT 2

AK049200

LOCUS

DEFINITION

AK049200 1841 bp mRNA linear HTC 20-SEP-2003

Mus musculus ES cells cDNA, RIKEN full-length enriched library,

clone: C330012K12 product: DUAL SPECIFICITY PROTEIN KINASE TTK (EC

2.7.1.-) (ESK) (PYT) homolog [Mus musculus], full insert sequence.

ACCESSION

AK049200

VERSION

AK049200.1 GI:26339883

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

REFERENCE

AUTHORS

Carninci, P. and Hayashizaki, Y.

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 1841)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/. Location/Qualifiers

1. 1841

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:C330012K12"

/db_xref="MGI:2416445"

/db_xref="taxon:10090"

/clone="C330012K12"

/cell_type="ES cells"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

85. >1839

/note="unnamed protein product; DUAL SPECIFICITY PROTEIN KINASE TTK (EC 2.7.1.-) (ESK) (PYT) homolog [Mus musculus] (SWISSPROT|P35761, evidence: FASTY, 99.9%ID, 96.7%length, match=2409) putative" /codon_start=1 /protein_id="BAC333605.1" /db_xref="GI:26339884" /translation="MEAEELIGSSVTIDSIMSKMRDIKNKINEDCTDELSLKICADH TETVNQIMRVGNTPEWNLNFKLLKXNSPLNDDLLNKLIGYSQAIEVLPDPKYGQN ESFARIQVRLAELKAIQEPDDARDYQFQVARENCCKFAFVHVSFAQFELSQGNLKSEQ LLHKAVETGAVPLQMLETAMRNHLQKQLLPEEDKKSVASVLSAQEPFSSSLGNV QNRSISCESRGQAARVLYGENLPQDAEVRHQNPFKQTHAAKRSCFPFGKVPVNLN SPDFYVKTDSAVTQLTRKSGSPDRDAILPGRSPRGSDSYELRGLKPIQTIYLDKSL VSNKSSSELMSDLIALKSKTDSLSLTLEETKPEIAERRPMQWQSTRKPECVFQNPAAF APLRHVPDVTPKADKESPPISVPKWLDPKSACETPSSSSSLDDYMKCFKTPVVKNDPPP ACPSSSTPYSQARLQOQQQGLSTPLQSLQISGSSSINECISVNGRIYSILKQIGSGG SSKVFQVLNEXKQINAIKYVNLDEADSQTIESYRNEIAFLNKLQOHSDKIIRLYDYEI TEQYIYVMVECGNIDLSW"

ORIGIN

Query Match 27.1%; Score 1047.2; DB 11; Length 1841; Best Local Similarity 76.5%; Pred. No. 3.1e-211; Matches 1410; Conservative 0; Mismatches 353; Indels 81; Gaps 7;

QY 971 TGTAGAAATGGAATCCGAGGATTTAAAGTGGCAGAGAAATTGACAATTGATTCATAATGAA 1030
Db 78 TTTAGAAATGGAGGCTGAAGAGTTAATTGGCAGCAGTGTGACGATTGATTCATCATGAG 137

QY 1031 CAAAGTGAGAGACATTTAAAATAAAGTTTAAAATAAAGACCTTACTGATGAAGCTT 1090
Db 138 CAAATGAGAGATATTAAAATAAG--ATAAATGAAGACTGTACTGATGAGCTAAGCTT 194

QY 1091 GAATAAAATTTCTGCTGATACAGATAACTCGGAACTGTTAACCAAAATTTATGATGAT 1150
Db 195 GTCTAAATCTGTG-----CCGATCACACGGAACCTGTTAACCAAAATTTATGAGGCT 245

QY 1151 GGCAAAACACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTC 1210
Db 246 TGGGAACACCCAGAGAGACTGGTTGAATTTTGTGCTGAACTAGAGAAAAACAGCTCACC 305

QY 1211 GCTAAGTGATGCTCTTTTAAATAAATTGATTGCTGTTTACAGTCAAGCAATTTGAAGCGCT 1270
Db 306 TCTAATGACGATCTTTTAAATAAAGTGTGTTGGTATAGTCAAGCGATTGAAGTACT 365

QY 1271 TCCCCCAGATAAATATGGCAAAATGAGAGTTTGTGCTAGAAATTCAGTGAGATTGCTGA 1330
Db 366 TCCTCCAGATAAATACGGCCAGAAATGAGAGCTTGTGCTGAATACAGTGAGACTTGTCTGA 425

QY 1331 ATTTAAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTTCAATGGCCAGAGCAA 1390
Db 426 ACTAAAGCTATTCAAGAGCCTGATGATGCCCGTACTACTTCCAGATGGCCAGGAA 485

QY 1391 CTGCAAGAAATTTGCTTTGCTCATATATCTTTGCACAAATTTGAACTGTCTACAAGGTAA 1450
Db 486 CTGCAAGAAATTTGCTTTGCTCACGATCTCTTTGCAAGTGTGAACTGTCTCAAGGCAA 545

QY 1451 TGTCAAAAAAGTAAACAACTTTCAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGTA 1510
Db 546 TCTTAAAAAAGTGAAGAGCTTCTTCAATAAAGCTGTAGAGACTGGGGCGGTGCCGCTGCA 605

QY 1511 AATGCTGGAATTTGCCCTGGGAATTTAAACCTCCAAAAAAGCAGTGTCTTTCAGAGGA 1570
Db 606 GATGCTGGAGACGGCCATGCGTAACTTACACCTCCAGAAAAAGCAGTGTCTTCCGAGGA 665

QY 1571 GGAAGAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGATCATTTTCCGGTTC 1630
Db 666 GGACAAAGAGAGTGTGTGAGCATCGACAGTACTAAGTGTCCCAAGAGCGGTTCTCCAGCTC 725

QY 1631 ACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGC 1690
Db 726 ACTTGGAAATGTACAGAAATCGGAGCATCAGCTGTGAGTCCAGAGGACAGGCTGGGGCAGC 785

QY	1691	CAGGTTTTATATGAGAGAAACATGCCACCAAGATGCAGAAATAGGTTACCGGAATTC	1750
Db	786	CAGGGTTTTATATGAGAGAACTTGCCCTCCACAAGATGCCGAAGTGAGGCATCAAAAACCC	845
QY	1751	ATTGAGACAAAATAACAAAACTAAACAGTCATGCCATTTGGAAGAGTCCAGTTAAACCT	1810
Db	846	CTTCAAGCAGACTCACGCAGCTAAACGGTCATGCCCTTTGGAAAAGTCCAGTCAATCT	905
QY	1811	TCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTATGAA	1870
Db	906	TCTAAACAGCCCGAGTTCTATGTGAAGACAGATAGCTCAGCTGTGACACACAGTTAAACAAC	965
QY	1871	AAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGG	1930
Db	966	AAGAAAGGGGTGAGACCAGA--CCGAGACGGGATTTGCCCCGCTCCAGACCCAGTGG	1022
QY	1931	AAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATTTCAAGGAACC	1990
Db	1023	CAGTGATTCCTATGAACCTGAGAGGTTTAAAGCCCATTCAAACTATCTATTTGAAAGACTC	1082
QY	1991	TCTGGTGCAGATGAAAAGAGTTCTGAACTTTATTAATTAATGATTTCAATAACCTGAAGAA	2050
Db	1083	TTTGGTGTCCAATGAAAAGAGTTCTGAAC--TTATGTCTGATTTAATAGCCTTGAAGAG	1139
QY	2051	TAAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATCAAGAACCCAGA	2110
Db	1140	TAAACAGATTTCAAG--TCTAACAAAATTTGGAAGAACTA-----AGCCAGA	1184
QY	2111	GGTTCAGAGAGTAACACAGAAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGTGTATTAAACCA	2170
Db	1185	GATTGCAGAAAAGAGGCCCATGCAAGTGGCAGTCTACCAGAAAAGCCGAGTGTGTGTTCCA	1244
QY	2171	GAATCCTGCTGATCTTCAAATCACTGGCAGATTCGGGAGTTAGCCCGAAAAGTTAATAC	2230
Db	1245	GAACCTGCTGCCCTTGGACCCCTGCGGCACGTTCCAGATGTACCCCGAAGGCTGA---	1301
QY	2231	AGAGCAGAAAACATACCACCTTTTGAGCAAACCTGTCTTTTTCAGTTTCAAACAGTCACCACC	2290
Db	1302	-----CAAAGAGTCACCACC	1316
QY	2291	AATATCAACATCTAAATGGTTTGACCCAAAATCTATTTGTAAGACACCAAGCAGCAATAC	2350
Db	1317	AATATCAGTTCCTAAAATGGCTTGATCCAAAAGTCTGCTGTGAGACACCTAGTAGCAGCTC	1376
QY	2351	CTTGATGATTACATGAGCTGTTTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGC	2410
Db	1377	CTTGGATGATTACATGAAATGTTTTTAAGACTCCAGTCCAGTTGTAAGAAATGACTTTCCACCTGC	1436
QY	2411	TTGTGAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAACAGCATCAAAT	2470
Db	1437	CTGTCCATCATCAACACCTTACAGCCAGCTTGCCCGCTCCAGCAGCAACAGCAGCAGGG	1496
QY	2471	ACTTGCCACTCCACTTCAAAAATTTACAGGTTTTAGCATCTTCTTTCAGCAAAATGAATGCAT	2530
Db	1497	ACTCAGCACTCCTCTTCAAAGCTTGCAGATTTTCAGGTTCTTCATCAATAATGAATGCAT	1556
QY	2531	TTCCGTTAAAGGAGAAATTTATTCCATATTTAAAGCAGATAGGAAGTGGAGTTCAAGCAA	2590
Db	1557	TTCAAGTTAACGGAAGAAATTTATTCCATATTTAAAGCAGATAGGCAGTGGAGTTCCAGTAA	1616
QY	2591	GGTATTTTCAGGTGTTAAATGAAAAGAAAACAGATATATGCTATAAAATATGTGAACTTAGA	2650
Db	1617	GGTGTTCAGGTATTGAATGAGAAAACAGATAAAGCTATCAAAATATGTGAACTTAGA	1676
QY	2651	AGAAGCAGATAACCAAACTCTTGATAGTTACCGGAACGAAAATAGCTTATTTGAAATAAACT	2710
Db	1677	AGACGCCGATAGCCAAAACCTATTGAGAGCTACCGCAACGAGATAGCGTTTTTTGAAACAACT	1736
QY	2711	ACAACAACACAGTGAAGATCATCCGACTTTATGATTAAGAAATCAAGGACCAAGTACAT	2770
Db	1737	ACAGCAACACAGTGAAGATCATCCGCTCTATGATTAAGAAATCAACGAGCAGTACAT	1796
QY	2771	CTACATGGTAATGGAGTGTGGAAAATATTGATCTTAATAGTTGGC	2814

||||| 1797 CTACATGGTAATGTAATGTGGAACATGACCTAAATAGTTGGC 1840

RESULT 3
BX415805
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX415805 1002 bp mRNA linear EST 13-MAY-2003
BX415805 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YE07
5-PRIME, mRNA sequence.
BX415805
BX415805.1 GI:30634668
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1002)
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3657.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP008AC04QP1&cluster=3657.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0CAP008AC04QP1.
Location/Qualifiers
1.1002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008YE07"
/tissue type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dt) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

FEATURES
source

ORIGIN

	Query Match	23.4%;	Score 906.2;	DB 13;	Length 1002;
	Best Local Similarity	98.8%;	Pred. No. 2.5e-181;		
	Matches 916;	Conservative	6;	Mismatches 4;	Indels 1;
QY	1999	CAGATGAAAAGAGTTC	TGAAC	TTATTATTACTGATT	CAATAACCCCTGAAGAAATAAAACGG 2058
Db	70	CGGATGAAAAGAGTTC	TGAAC	TTATTATTACTGATT	CAATAACCCCTGAAGAAATAAAACGG 129
QY	2059	AATCAAGTCTTCTAG	CTAAAT	TAGAAGAAACTAA	AGAGTATCAAGAACCCAGAGGTTCCAG 2118
Db	130	AATCAAGTCTTCTAG	CTAAAT	TAGAAGAAACTAA	AGAGTATCAAGAACCCAGAGGTTCCAG 189
QY	2119	AGAGTAACCCAGAAA	CAGTGG	CAAGCTAAAGAGAA	AGTCCAGAGTGATTAACCCAGAATCCTG 2178
Db	190	AGAGTAACCCAGAAA	CAGTGG	CAATCTAAGAGAA	AGTCCAGAGTGATTAACCCAGAATCCTG 249
QY	2179	CTGCATCTTCAAAT	CAC	TGGCAGATTCGG	GAGTTAGCCCGAAAAGTTAATACAGAGCAGA 2238
Db	250	CTGCATCTTCAAAT	CAC	TGGCAGATTCGG	GAGTTAGCCCGAAAAGTTAATACAGAGCAGA 309
QY	2239	AACATACCAC	TTTTGAG	CAACCTGCTCTTT	TCAGTTTCAAAACAGTCACCCACCAATATCAA 2298
Db	310	AACATACCAC	TTTTGAG	CAACCTGCTCTTT	TCAGTTTCAAAACAGTCACCCACCAATATCAA 369
QY	2299	CATCTAAATG	TTGAC	CCCAAAATCTATT	TGTAAGACACCAAGCAGCAATACCTTGGATG 2358

Db 370 CATCTAAATGGTTTGACCCAAATCTATTGTGAAGACACCAAGCAGCAATACCTTGGATG 429

QY 2359 ATTACATGAGCTGTTTITAGAACTCCAGTTGTAAAGATGACTTTCCACCTGCTTGTTCAGT 2418

Db 430 ATTACATGAGCTGTTTITAGAACTCCAGTTGTAAAGATGACTTTCCACCTGCTTGTTCAGT 489

QY 2419 TGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAACAGCATCAATACTTGGCCA 2478

Db 490 TGTCAACACCTTATGGCCAACTGCTGTTTCCAGCAGCAATAGCATCAATACTTGGCCA 549

QY 2479 CTCCACTTCAAAATTTACAGGTTTITAGCATCTTCTTACGCAAAATGAATGATTCGGTTA 2538

Db 550 CTCCACTTCAAAATTTACAGGTTTITAGCATCTTCTTACGCAAAATGAATGATTCGGTTA 609

QY 2539 AAGGAAGATTTATTCATATTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTC 2598

Db 610 AAGGAAGATTTATTCATATTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTC 669

QY 2599 AGGTGTTAAATGAAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGAACGAG 2658

Db 670 AGGTGTTAAATGAAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGAACGAG 729

QY 2659 ATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACAAC 2718

Db 730 ATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATTTGAATAAACTACAACAAC 789

QY 2719 ACAGTGATAAGATCATCCGACTTTATGATCTTAATAGTTGGCTTAAAAAGAAAAATCCATTGATC 2778

Db 790 ACAGTGATAAGATCATCCGACTTTATGATCTTAATAGTTGGCTTAAAAAGAAAAATCCATTGATC 849

QY 2779 TAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAAAAGAAAAATCCATTGATC 2838

Db 850 TAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAAAAGAAAAATCCATTGATC 909

QY 2839 CATGGGAACGCAAGAGTTACTGGAAGAAATATGTTAGAGCAGTTTACACACATCCATCAAC 2898

Db 910 CATGGGAACGCAAGAGTTACTGGAAGAAATATGTTAGAGCAG-TCMCACATCCATCAAC 968

QY 2899 ATGGCAATGTTACAGTGATCTTAAAC 2925

Db 969 ATGGCAATGTTACAGTGATCTTAAAC 995

RESULT 4

AU132784

LOCUS

DEFINITION AU132784 NT2RP4 Homo sapiens cDNA clone NT2RP4000532 5', mRNA

sequence.

ACCESSION AU132784

VERSION AU132784.1 GI:10993323

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 887)

AUTHORS Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and Isogai,T.

TITLE HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y., Isogai,T.)

Unpublished (2000)

CONTACT: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and

FEATURES

source

Location/Qualifiers

1..887

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP4000532"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/clone_lib="NT2RP4"

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Query Match 22.0%; Score 851.8; DB 9; Length 887;

Best Local Similarity 98.8%; Pred. No. 8.5e-170;

Matches 878; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 1105 CTGATACACAGATAAAGCTGGGAAGTGTAAACCAATTTATGATGATGGCAACACCCAG 1164

Db 1 CTGATACACAGATAAAGCTGGGAAGTGTAAACCAATTTATGATGATGGCAACACCCAG 60

QY 1165 AGGACTGGTTCAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCGCTAAGTATGCTC 1224

Db 61 AGGACTGGTTCAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCGCTAAGTATGCTC 120

QY 1225 TTTTAAATAAATTTGATTGGTTCAGTCAAGCAATTTGAAGCGCTTCCCCAGATAAAT 1284

Db 121 TTTTAAATAAATTTGATTGGTTCAGTCAAGCAATTTGAAGCGCTTCCCCAGATAAAT 180

QY 1285 ATGGCCAAATGAGAGTTTGTGCTAGAATTTCAAGTGAAGTTTCTGAATTTAAAGCTATT 1344

Db 181 ATGGCCAAATGAGAGTTTGTGCTAGAATTTCAAGTGAAGTTTCTGAATTTAAAGCTATT 240

QY 1345 AAGAGCCAGATGATGCACGTGACTACTTTCAAATGGCCAGAGCAAACTGCAAGAAAATTG 1404

Db 241 AAGAGCCAGATGATGCACGTGACTACTTTCAAATGGCCAGAGCAAACTGCAAGAAAATTG 300

QY 1405 CTTTGTGTCATATATCTTTTGCACAATTTTGAAGTGTCAAGGTAATGTCAAAAAAAGTA 1464

Db 301 CTTTGTGTCATATATCTTTTGCACAATTTTGAAGTGTCAAGGTAATGTCAAAAAAAGTA 360

QY 1465 AACAACTTCTTCAAAAAGCTGTAGAACGCTGGAGCAGTACCAGTAAATGCTGGAATTTG 1524

Db 361 AACAACTTCTTCAAAAAGCTGTAGAACGCTGGAGCAGTACCAGTAAATGCTGGAATTTG 420

QY 1525 CCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGTTTTCAGAGGAGGAAAAAGAAATT 1584

Db 421 CCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGTTTTCAGAGGAGGAAAAAGAAATT 480

QY 1585 TATCAGCATCTACGGTATTAACTGCCCCAAGAAATCATTTTCCGGTTCACITGGGCATTTAC 1644

Db 481 TATCAGCATCTACGGTATTAACTGCCCCAAGAAATCATTTTCCGGTTCACITGGGCATTTAC 540

QY 1645 AGAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTTATATG 1704

Db 541 AGAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTAAAGCCAGGTTTATATG 600

QY 1705 GAGAGAACATGCCACCACCAAGATGCAGAAAATAGTTTACCGGAATTCATTGAGACAACTA 1764

Db 601 GAGAGAACATGCCACCACCAAGATGCAGAAAATAGTTTACCGGAATTCATTGAGACAACTA 660

QY 1765 ACAAACTAAACAGTCAATGCCATTTTGGAGAGTCCAGTTAACTTCTAAATAGCCCCAG 1824

Db 661 ACAAACTAAACAGTCAATGCCATTTTGGAGAGTCCAGTTAACTTCTAAATAGCCCCAG 720

QY 1825 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATTTATGAAAAAGACAACTCTA 1884

Db 721 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATTTATGAAAAAGACAACTCTA 779

QY 1885 GATCAGATGCCGAGATTTGTTGTGCTGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 1944

Db 780 GATCAGATGCCGAGATTTGTTGTGCTGCTGGATCTAAACCAAGTGGAAATGATTCCTGTG 839

QY	1267	CGCTTCCCCAGATAAAATATGGCCAAAATGAGAGTTTTTGCTAGAAATCAAGTGAGATTG	1326
Db	301	CGCTTCCCCAGATAAAATATGGCCAAAATGAGAGTTTTTGCTAGAAATCAAGTGAGATTG	360
QY	1327	CTGAATTAAGAAGCTATTCAAGAGGCCAGATGCACGTGACTACTTTCAAATGGCCAGAG	1386
Db	361	CTGAATTAAGAAGCTATTCAAGAGGCCAGATGCACGTGACTACTTTCAAATGGCCAGAG	420
QY	1387	CAAACTGCAAGAAAAATTTGCTTTTGTTCATATATCTTTTGCAACAATTTGAACTGTCAACAAG	1446
Db	421	CAAACTGCAAGAAAAATTTGCTTTTGTTCATATATCTTTTGCAACAATTTGAACTGTCAACAAG	480
QY	1447	GTAATGTCAAAAAAAGTAAACAACCTTCTTCAAAAAAGCTGTAGAACGTGGAGCAGTACCAC	1506
Db	481	GTAATGTCAAAAAAAGTAAACAACCTTCTTCAAAAAAGCTGTAGAACGTGGAGCAGTACCAC	540
QY	1507	TAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAACCTTCAAAAAAGCAGCTGCTTTTCAG	1566
Db	541	TAGAAATGCTGGAAATTTGCCCTGCGGAATTTANACCTCCAACCTTCAAAAAAGCAGCTGCTTTTCAG	600
QY	1567	AGGAGGAAAAGAAGAAATTTATCAGCATCTACGGTATTAACTGCCCCAAGAATCATTTTCCG	1626
Db	601	AGGAGGAAAAGAAGAAATTTATCAGCATCTACGGTATTAACTGCCCCAAGAATCATTTTCCG	660
QY	1627	GTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTA	1686
Db	661	GTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTTCCAGAGGACAGACTACTA	720
QY	1687	AAGCCAGGTTTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTTACCGGA	1746
Db	721	AAGCCAGGTTTTTATATGGAGAGAACATGCCACCACAAGATGCAGAAATAGGTTACCGGA	780
QY	1747	ATTCAATTGAGACAAAACCTAACAAAACCTAACAGTCATGCCCATTTTGGAGAGTCCCAGTTA	1806
Db	781	ATTCAATTGAGACAAAACCTAACAAAACCTAACAGTCATGCCCATTTTGGAGAGTCCCAGTTA	840
QY	1807	ACC 1809	
Db	841	AAC 843	
RESULT	6		
LOCUS	AU124386		
DEFINITION	AU124386 NT2RM2 Homo sapiens cDNA clone NT2RM2002111 5', mRNA sequence.		
ACCESSION	AU124386		
VERSION	AU124386.1 GI:10949102		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 916)		
	Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and Isogai, T.		
TITLE	HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S., Isogai, T.)		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers		
FEATURES			

FEATURES

```
source 1. .916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM2002111"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RM2"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"

ORIGIN
Query Match 21.0%; Score 812.8; DB 9; Length 916;
Best Local Similarity 95.7%; Pred. No. 1.5e-161;
Matches 877; Conservative 0; Mismatches 31; Indels 8; Gaps 4;

QY 1649 TAGGAACAACAGTGTGATCCAGAGGACAGACTACTAAAGCCAGGTTTATATGGAGA 1708
Db 1 TAGGAACAACAGTGTGATCCAGAGGACAGACTACTAAAGCCAGGTTTATATGGAGA 60
QY 1709 GAACATGCCACCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGACAAACTAACAA 1768
Db 61 GAACATGCCACCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGACAAACTAACAA 120
QY 1769 AACTAAACAGTCATGCCCATTTGGAAGAGTCCCAGTTAACCTTCTAAATAGCCAGATTG 1828
Db 121 AACTAAACAGTCATGCCCATTTGGAAGAGTCCCAGTTAACCTTCTAAATAGCCAGATTG 180
QY 1829 TGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGAAAGACAAACCTCTAGATC 1888
Db 181 TGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGAAAGACAAACCTCTAGATC 240
QY 1889 AGAATGCCGAGATTGGTTGTGCTGGATCTAAACCAAGTGGAATGATTCCTGTGAATT 1948
Db 241 AGAATGCCGAGATTGGTTGTGCTGGATCTAAACCAAGTGGAATGATTCCTGTGAATT 300
QY 1949 AAGAAATTTAAAGTCTGTTCAAAATAGTCATTTCAAGGAACCTCTGGTGCAGATGAAAA 2008
Db 301 AAGAAATTTAAAGTCTGTTCAAAATAGTCATTTCAAGGAACCTCTGGTGCAGATGAAAA 360
QY 2009 GAGTTCGAACTTATTATTACTGATTCATAATAACCTGAAGATAAAACGGAATCAAGTCT 2068
Db 361 GAGTTCGAACTTATTATTACTGATTCATAATAACCTGAAGATAAAACGGAATCAAGTCT 420
QY 2069 TCTAGCTAAATTAGAAGAAACTAAAGAGTATCAAGAACCCAGAGTTCCAGAGAGTAACCA 2128
Db 421 TCTAGCTAAATTAGAAGAAACTAAAGAGTATCAAGAACCCAGAGTTCCAGAGAGTAACCA 480
QY 2129 GAAACAGTGGCAAGCTAAGAAAGTCAGAGTGATTAACCCAGAAATCCTGCTGCATCTTC 2188
Db 481 GAAACAGTGGCAATCTAAGAAAGTCAGAGTGATTAACCCAGAAATCCTGCTGCATCTTC 540
QY 2189 AAATCACTGGCAGATTCCGGAGTTAGCCCGAAAGTTAATACAGAGCAGAAACATACCAC 2248
Db 541 AAATCACTGGCAGATTCCGGAGTTAGCCCGAAAGTTAATACAGAGCAGAAACATACCAC 600
QY 2249 TTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCACCAACCAATATCAACATCTAAATG 2308
Db 601 TTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCACCAACCAATATCAACATCTAAATG 660
QY 2309 GTTTGACCCAAAATCTATTGTAAGACACCAAGCAGCAATACCTTGGATGATTACATGAG 2368
Db 661 GTTTGACCCAAAATCTATTGTAAGACACCAAGCAGCAATACCTTGGATGATTACATGAG 720
QY 2369 CTGTTTGTAGAACTCCAGTTGT- AAAGAATGACTTTCCACCTGCTTGTGTCAGTTGTCAACAC 2427
Db 721 CTGTTTGTAGAACTCCAGTTGTGTAAGAAGATGACTTTCACCTGCTTGTGTCAGTTGTCAACAC 779
QY 2428 CTTATGCGCAACCTGCCTGTTTCCAGCAGCAACAGCATCAAAATACTTGCCTACT- CCACT 2485
Db 780 CTTATGCGCAACCTGCCTGTTTCCAGCAGCAACAGCATCAAAATACTTGCCTACTTCCACTT 839
QY 2486 TCAAAATTTACAGGTTTTCAGCATCTTCTTCAGCAAA- - - -TGAATGCATTTTCGGTTAAAG 2541
```

```
Db 840 CAAAAATTACAGGTTTAAACCTCTTCTTTAACCAATGGAATGCCCTTCNGGTTAANG 899
QY 2542 GAAGAAATTTATTCAT 2557
Db 900 GACGAATTTTNTTCT 915

RESULT 7
BQ929926 931 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_8950097 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6471444
DEFINITION 5', mRNA sequence.
ACCESSION BQ929926
VERSION BQ929926.1 GI:22344957
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 931)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14003 row: m column: 13
High quality sequence stop: 656.
FEATURES
Location/Qualifiers
1..931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6471444"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."

ORIGIN
Query Match 21.0%; Score 811.4; DB 13; Length 931;
Best Local Similarity 94.7%; Pred. No. 3e-161;
Matches 869; Conservative 0; Mismatches 40; Indels 9; Gaps 3;

QY 2640 GTGAACCTTAGAAGAGCAGATAACCAAACTCTTGATAGTTACCGAACGAAATAGCTTAT 2699
Db 1 GTGAACCTTAGAAGAGCAGATAACCAAACTCTTGATAGTTACCGAACGAAATAGCTTAT 60
QY 2700 TTGAATAAACTACAACAACACAGTGATAAGATCATCCGACTTTTATGATTAATAATCAG 2759
Db 61 TTGAATAAACTACAACAACACAGTGATAAGATCATCCGACTTTTATGATTAATAATCAG 120
QY 2760 GACCAATACATCTACATGTTAATGGAGTGTGGAATAATTGATCTTAATAGTTGGCTTAA 2819
Db 121 GACCAATACATCTACATGTTAATGGAGTGTGGAATAATTGATCTTAATAGTTGGCTTAA 180
QY 2820 AAGAAAAAATCCATTGATCCATGGGAACGGAAGAGTTACTGGAAAAATATGTTAGAGGCA 2879
Db 181 AAGAAAAAATCCATTGATCCATGGGAACGGAAGAGTTACTGGAAAAATATGTTAGAGGCA 240
QY 2880 GTTCAACAATCCATCAACATGGCATTGTTACAGTGATCTTAAACCAGCTAATCTTCTG 2939
Db 241 GTTCAACAATCCATCAACATGGCATTGTTACAGTGATCTTAAACCAGCTAATCTTCTG 300
```

QY 2940 ATAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGAT 2999
Db |||||||
QY 301 ATAGTTGATGGAATGCTAAAGCTAATTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGAT 360
Db |||||||
QY 3000 ACAACAAGTGTGTTAAAGATTCTCAGGTTGGCACAGTTAATTATATATGCCCACCAAGACA 3059
Db |||||||
QY 361 ACAACAAGTGTGTTAAAGATTCTCAGGTTGGCACAGTTAATTATATATGCCCACCAAGACA 420
QY 3060 ATCAAAGATATGTCCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAA 3119
Db |||||||
QY 421 ATCAAAGATATGTCCTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAA 480
QY 3120 AGTGATGTTTGGTCCTTAGGATGTAATTTTGTACTATATGACTTACGGGAAAACACCAATTT 3179
Db |||||||
QY 481 AGTGATGTTTGGTCCTTAGGATGTAATTTTGTACTATATGACTTACGGGAAAACACCAATTT 540
QY 3180 CAGCAGATAATTAATCAGATTTCTAAATTAACATGCCATAATTTGATCCTAATCATGAATTT 3239
Db |||||||
QY 541 CAGCAGATAATTAATCAGATTTCTAAATTAACATGCCATAATTTGATCCTAATCATGAATTT 600
QY 3240 GAATTTCCCGATATTCAGAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGG 3299
Db |||||||
QY 601 GAATTTCCCGATATTCAGAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGG 660
QY 3300 GACCCAAAACAGAGGATATCCATTCCTGAGTCCTGGCTCATCCATATGTTCAAATTCAA 3359
Db |||||||
QY 661 GACCCAAAACAGAGGATATCCATTCCTGAGTCCTGGCTCATCCATATGTTCCNNAATCA 720
QY 3360 ACTCATCCAGTTAACCAAAATGCCAAGGGGAACCACTGAAGAATGAATATGTTCTGGGC 3419
Db |||||||
QY 721 ACTCATCCAGTTAACCAAAATGCCAAGGGGAACCACTGAAGAANTGAATATGTTCTGGC 780
QY 3420 CAACTTGTGCTGTAATTCCTTAACCTCAATTTTGAAAGCTGCTAAACCTTTATATGA - 3478
Db |||||||
QY 781 CAACTTGTGCTGTAATTCCTTAACCTCAATTTTGAAAGCTGCTAAACCTTTATATGA 840
QY 3479 --ACACTATAGTGGTGAAAGTCATAATT--CTTCATCCTCC---AAGACTTTTGAA 3530
Db |||||||
QY 841 ACACTATAGTGGTGAAAGTCATAATTCCTTAACCTCCACAGGAACCTTTTITGAA 900
QY 3531 AAAAAAAGGGGAAAAAAA 3548
Db |||||||
QY 901 AAAAAAGGGGAAAAAAA 918

RESULT 8
BX361758/c
LOCUS BX361758 1059 bp mRNA linear EST 05-MAY-2003
DEFINITION BX361758 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ004YA24 3-PRIME, mRNA sequence.
ACCESSION BX361758
VERSION BX361758.1 GI:30376559
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1059)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3657.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ004BA12NP1&cluster=3657.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ004BA12NP1.

FEATURES
source Location/Qualifiers
1..1059
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ004YA24"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 20.3%; Score 783.8; DB 13; Length 1059;
Best Local Similarity 93.9%; Pred. No. 2.1e-155;
Matches 820; Conservative 24; Mismatches 24; Indels 5; Gaps 3;
QY 2950 GAATGCTAAAGCTAATTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGATCAACAAGTG 3009
Db |||||||
QY 871 GAATGCTAAAGCTAATKGA-TTTGGGATGCCAAVCCAAGGCARCCAGATACAACAAGTG 813
QY 3010 TTGTAAAGATTCTCAGGTTGGCACAGTTAATTATATATGCCACCAGAAAGCAATCAAGATA 3069
Db |||||||
QY 812 TTGTAAAGAKTCTCAGGTTGGCACAGTTAATTATATATGCCACCAGADGCAWTCAAGATA 753
QY 3070 TGTCTTCTCCAGAGAGAAATGGAATCTAAGTCAAAAGATAAGCCCCAAAAGTGATGTTT 3129
Db |||||||
QY 752 TGTCTTCTCCAGAGAGAAATGGAATCTAAGTCAAAAGATAAGCCCCAAAAGTGATGTTT 693
QY 3130 GGTCCTTAGGATGTAATTTTGTACTATATGACTTACGGGAAAACACCATTTTCAGCAGATAA 3189
Db |||||||
QY 692 GGTCCTTAGGATGTAATTTTGTACTATATGACTTACGGGAAAACACCATTTTCAGCAGATAA 633
QY 3190 TTAATCAGATTTCTAAATTACATGCCATAATTGATCCTAATCATGAATTTGAATTTCCCG 3249
Db |||||||
QY 632 TTAATCAGATTTCTAAATTACATGCCATAATTGATCCTAATCATGAATTTGAATTTCCCG 573
QY 3250 ATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAAAC 3309
Db |||||||
QY 572 ATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAAAC 513
QY 3310 AGAGATATCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAACTCATCCAG 3369
Db |||||||
QY 512 AGAGATATCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAACTCATCCAG 453
QY 3370 TTAACCAAAATGGCCAAGGAAACCACTGAAGAATGAAATATGTTCTGGGCCAACTTGTG 3429
Db |||||||
QY 452 TTAACCAAAATGGCCAAGGAAACCACTGAAGAATGAAATATGTTCTGGGCCAACTTGTG 393
QY 3430 GTCTGAATTTCTCCTAATCCATTTTGAAAGCTGCTAAACCTTTATATGAACACTATAGTG 3489
Db |||||||
QY 392 GTCTGAATTTCTCCTAATCCATTTTGAAAGCTGCTAAACCTTTATATGAACACTATAGTG 333
QY 3490 GTGTTGAAAGTCATAATTTCTCATCCTCCAGACTTTTGAAAAAAGGGGAAAAAAT 3549
Db |||||||
QY 332 GTGTTGAAAGTCATAATTTCTCATCCTCCAGACTTTTGAAAAAAGGGGAAAAAAT 273
QY 3550 GATTTGCAGTTATTTCGTAATGTCAGATAGGAGGTATAAAATATA---TTGGACTGTTATA 3606
Db |||||||
QY 272 GATTTGCAGTTATTTCGTAATGTCAGATAGGAGGTATAAAATATA---TTGGACTGTTATA 213
QY 3607 CTCTTGAATCCCTGTGGAAATCTACATTTGAAGACCAACATCCTCTGAAGTGTATCAGC 3666
Db |||||||
QY 212 CYCTTSAATCCCTGTGGAAATCTACATTTGAARACACATCCTCTGAAGTGTATCAGC 153
QY 3667 AAAAAAATTCAGTGAGATTATCTTTTAAAGAAAACTGTAAAAATAGCAACCACTTATGG 3726
Db |||||||
QY 152 AAAAAAATTCAGT-AGATTATCTTTTAAAGAAAACTGTAAAAATAGCAACVACTTATGG 94
QY 3727 CACTGTATATATTGTAGACTTGTGTTTCTCTCTGTTTATGCTCTTGTGTAATCTACTTGACA 3786


```
Db      93  KACTGATWTTTKWKAYTKTNTTTTATGTTTCTGTAATTACTTGACW 34
QY      3787 TCATTTTACTCTTGAATAGTGGTGGATAGCA 3819
Db      33  TYTTTTTATTCTTGAATAGTGGTGGATAGCA 1

RESULT 9
AU133100
LOCUS
DEFINITION AU133100 NT2RP4 Homo sapiens cDNA clone NT2RP4001283 5', mRNA
sequence.
ACCESSION AU133100
VERSION AU133100
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and
Isogai,T.
TITLE HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T.,
Sugano,S., Masuho,Y., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP4001283"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RP4"
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

ORIGIN
Query Match 20.1%; Score 777.4; DB 9; Length 920;
Best Local Similarity 97.8%; Pred. No. 4.8e-154;
Matches 840; Conservative 0; Mismatches 13; Indels 6; Gaps 5;
QY 1150 TGGCAACACACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTC 1209
Db 1 TGGCAACACACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTC 60
QY 1210 CGCTAAGTAGTCTCTTTTAAATAAATTGATTGGTCTTACAGTCAAGCAATTGAAGCGC 1269
Db 61 CGCTAAGTAGTCTCTTTTAAATAAATTGATTGGTCTTACAGTCAAGCAATTGAAGCGC 120
QY 1270 TTCCCCCAGATAAAATATGGCCAAAATGAGAGTTTGTGCTAGAAATTCAGTGAGATTGCTG 1329
Db 121 TTCCCCCAGATAAAATATGGCCAAAATGAGAGTTTGTGCTAGAAATTCAGTGAGATTGCTG 180
QY 1330 AATTTAAAGCTATTCAAGAGCCAGATGTCACGTGACTACTTTTCAATGGCCAGAGCAA 1389
Db 181 AATTTAAAGCTATTCAAGAGCCAGATGTCACGTGACTACTTTTCAATGGCCAGAGCAA 240
QY 1390 ACTGCAAGAAATTGCTTTTGTTCATATATCTTTTGCACAATTGAACTGTCACAAGGTA 1449
```

```
Db      241 ACTGCAAGAAATTGCTTTTGTTCATATATCTTTTGCACAATTGAACTGTCAAGGTA 300
QY 1450 ATGTCAAAAAAGTAAACAACCTTCTTCAAAAAAGCTGTAGAACGCTGGAGCAGTACCCTAG 1509
Db 301 ATGTCAAAAAAGTAAACAACCTTCTTCAAAAAAGCTGTAGAACGCTGGAGCAGTACCCTAG 360
QY 1510 AAATGCTGGAAATTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTCTTTTCAGAGG 1569
Db 361 AAATGCTGGAAATTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTCTTTTCAGAGG 420
QY 1570 AGGAAAAGAAGAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCATTTTCCGGTT 1629
Db 421 AGGAAAAGAAGAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCATTTTCCGGTT 480
QY 1630 CACTTGGGCATTTACAGAATAGGAACAACAGTTTGATTCAGAGGACAGACTACTAAAG 1689
Db 481 CACTTGGGCATTTACAGAATAGGAACAACAGTTTGATTCAGAGGACAGACTACTAAAG 540
QY 1690 CCAGGTTTTTATATGGAGAGAACATGCCACCAAGATGCAGAAATAGTTTACCGGAATT 1749
Db 541 CCAGGTTTTTATATGGAGAGAACATGCCACCAAGATGCAGAAATAGTTTACCGGAATT 600
QY 1750 CATTGAGACAAAACCTAACAAAACCTAAACAGTCATGCCATTGGAAAGAGTCCCAAGTTAAC 1809
Db 601 CATTGAGACAAAACCTAACAAAACCTAAACAGTCATGCCATTGGAAAGAGTCCCAAGTTAAC 660
QY 1810 TTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTTATGA 1869
Db 661 TTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGT-CTTGGTTTATGA 719
QY 1870 AAAGACAAACCTCTAGATCAGAATG-CCGAGATTTGGTGTGCTGGATCTAAACCAAGT 1928
Db 720 AAAGACAAACCTCTAGATCAGAATGCCCCGAGATTTGGTGTGCTGGATCTAAACCAAGT 779
QY 1929 GGAATGATTCCTGTGAATTAAGAAATTTAAAGTGTGTTCAAAATAGTCAATTTCAAGGAA 1988
Db 780 GGAATGATTCCTGTGNGAATTAAGAAATTTAAAG-CTGGTCAAAATAG--CATTCAGGAA 836
QY 1989 CCTCTGGTGTGATGAAA 2007
Db 837 CCT-TGGTGTCAATGAAA 854

RESULT 10
AU119487
LOCUS
DEFINITION AU119487 HEMBA1 Homo sapiens cDNA clone HEMBA1005932 5', mRNA
sequence.
ACCESSION AU119487
VERSION AU119487.1 GI:10934722
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 837)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
```



```
Db 306 AGTCAAGCAATGAAGCGCTTCCCCAGATATAATATGGCCAAATGAGAGTTTGTCTAGA 365
QY 1311 ATTCAAGTGAGATTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGACGTACTAC 1370
Db 366 ATTCAAGTGAGATTGCTGAATTAAGAGCTATTCAAGAGCCAGATGATGACGTACTAC 425
QY 1371 TTTCAAAATGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAA 1430
Db 426 TTTCAAAATGCCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAA 485
QY 1431 TTTGAAGCTGTCACAAAGGTAATGTCAAAAAAAGTAACAACCTTTCTTCAAAAAGCTGTAGAA 1490
Db 486 TTTGAAGCTGTCACAAAGGTAATGTCAAAAAAAGTAACAACCTTTCTTCAAAAAGCTGTAGAA 545
QY 1491 CGTGGAGCAGTACCACACTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAA 1550
Db 546 CGTGGAGCAGTACCACACTAGAAATGCTGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAA 605
QY 1551 AAGCAGCTGCTTTC--AGAGGAGGAAAGAAAGTAATTTATCAGCATCTACGGTATTAACT 1607
Db 606 AAGCAGCTGCTTTCAGAGGAGGAAAGAAAGTAATTTATCAGCATCTACGGTATTAACT 665
QY 1608 GCCCAAGATCATTTTCCGGTTCACTTGGGCATTTTACAGATAGGAACACAGTTGTGAT 1667
Db 666 GCCCAAGATCATTTTCCGGTTCACTTGGGCA-TTACAGATAGGAACACAGTTGTGAT 724
QY 1668 TCCAGAGGACAGACTACTAAAGCCAGGTTTTTATATGAGAGAACATGCCACCACAAAGAT 1727
Db 725 TCCAGAGGACAGACTACTAAAGCCAGGTTCTCCATATGAGAGAACATGTCA-CACAAGAT 783
QY 1728 GCAGAAATAGCTTACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCA 1787
Db 784 GC-GAAATAGCTTACCGGAATTTCAATTGAGACAACTAAACAAACTAAACAGGTCATGCCA 842
QY 1788 TTTGGAAGTCCCAGTTAACTTCTAA 1815
Db 843 TTGGAAGATTCCCAGTTAACTTCTAA 870
```

```
RESULT 12
AA744546/c
LOCUS
DEFINITION
ny79c06.s1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:1284490 3'
similar to gb:M86699 DUAL SPECIFICITY PROTEIN KINASE TTK (HUMAN);,
mRNA sequence.
ACCESSION
AA744546
VERSION
AA744546.1 GI:2783310
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 912)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1024 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 479.
Location/Qualifiers
```

```
source
1. .912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1284490"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/clone_lib="NCI CGAP GCBI"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CHBR). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCTCATTTTTTTTTTTTTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

ORIGIN

Query Match	Best Local Similarity	Score	DB 9;	Length	912;
Matches	836;	Conservative	0;	Mismatches	53;
				Indels	7;
				Gaps	6;
QY 2971	TTGGGATTGCAACCAATGCAACCCAGATACACAAAGTGTGTTAAAGATTCTCAGGTTG	3030			
Db 899	TTTGGGTCTCAACCCACTGCCCCCGATCCAC-AGTGTGGTTAAGATTCTCTCGGTGG	841			
QY 3031	GCACAGTTAATTATATGCCACAGAGCAATCAAGATATGTCTTCTCCAGAGAGATG	3090			
Db 840	CCACGGTTAAT--TTAGCCCAAGAGACCAACAGGATATGTCTTCTC-CAGTGAGATG	784			
QY 3091	GGAAATCTAAGTCAAAGATAAGCCCCCAAGAGTGTGTTGGTCTTAGATGATTATTGT	3150			
Db 783	GGAACTCTAAGTC-AAGATAAGCCCCCAAGTGTGTTGTTCTTAGTATGATTATTGT	725			
QY 3151	ACTATATGACTTACGGGAAACACCATTTTCAGCAGATAATTAATCAGATTCTTAAATTAC	3210			
Db 724	ACTATATGACTTACGGGAAACACCATCTCAGCAGATAATAAATCAGATTCTTAAATTAC	665			
QY 3211	ATGCCATAATTGATCCTAATCATGAAATTGAATTTCCCGATATTTCCAGAGAAAGATCTTC	3270			
Db 664	ATGCCATAATTGATCCTAATCATGAAATTGAATTTCCCGATATTTCCGAGAAAGATCTTC	605			
QY 3271	AAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAACAGAGGATATCCATTCTCTGAGC	3330			
Db 604	AAGATGTGTTAAAGTG-TGTTTAAAGGGACCCAAACAGAGGATATCCATTCTCTGAGC	546			
QY 3331	TCCTGGCTCATCCATATGTTCAAATTCAACTCATCCAGTTAACCAATGGCCAAAGGAA	3390			
Db 545	TCCTGGCTCATCCATATGTTCAAATTCAACTCATCCAGTTAACCAATGGCCAAAGGAA	486			
QY 3391	CCACTGAAGAAATGAAATATGTTCTGGCCAACTTGTGTTGTTCTGAATTTCTTAACCTCA	3450			
Db 485	CCACTGAAGAAATGAAATATGTTCTGGCCAACTTGTGTTGTTCTGAATTTCTTAACCTCA	426			
QY 3451	TTTTGAAAGCTGCTAAAATTTATATGAACACTATAGTGGTGGTGAAGTCAATAATCTT	3510			
Db 425	TTTTGAAAGCTGCTAAAATTTATATGAACACTATAGTGGTGGTGAAGTCAATAATCTT	366			
QY 3511	CATCCTCAAGACTTTTGAAGAAAGGGGAAAGAAATGATTTCAGTTATTCGTAATG	3570			
Db 365	CATCCTCAAGACTTTTGAAGAAAGGGGAAAGAAATGATTTCAGTTATTCGTAATG	306			
QY 3571	TCAGATAGGAGGTATAAAATATATTGGACTGTTTACTCTTGAATCCCTGTGGAATCTA	3630			
Db 305	TCAGATAGGAGGTATAAAATATATTGGACTGTTTACTCTTGAATCCCTGTGGAATCTA	246			
QY 3631	CATTGAGACAAACATCACTCTGAAGTGTATATCAGCAAAAAAATTCAGTGAGATTATCT	3690			

Db 245 CATTGAAGACAATCACTCTGAAGTGTATATCAGCAAAAAAATTCAGT-AGATTATCT 187

Qy 3691 TTAAGAGAAACTGTAAAAATAGCAACCACTTATGGCACTGTATATATTGTAGACTTGT 3750

Db 186 TTAAGAGAAACTATAAAAATAGCAACCACTTATGGCACTGTATATATTGTAGACTTGT 127

Qy 3751 TTCTCTGTTTATGCTCTTGTGTAAATCTACTTGACATCATTTTACTCTTGGAAATAGTGGG 3810

Db 126 TTCTCTGTTTATGCTCTTGTGTAAATCTACTTGACATCATTTTACTCTTGGAAATAGTGGG 67

Qy 3811 TGGATAGCAAGTATATCTAAAAAACTTTGTAAATAAAGTTTGTGGCTAAAAATGA 3866

Db 66 TGGATAGCAAGTATATCTAAAAAACTTTGTAAATAAAGTTTGTGGCTAAAAATGA 11

RESULT 13

BQ428904

LOCUS

DEFINITION BQ428904 7905694 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6105546 5', mRNA sequence.

ACCESSION BQ428904

VERSION BQ428904.1 GI:21167980

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 792)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: CLONTECH cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Plate: LLCM2345 row: c column: 19 High quality sequence stop: 548.

FEATURES

source 1..792 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6105546" /lab_host="DH10B (T1 phage-resistant)" /clone_lib="NIH_MGC_82" /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccctggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGGCGGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 18.9%; Score 731.4; DB 13; Length 792;

Best Local Similarity 98.6%; Pred. No. 2.7e-144;

Matches 768; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

Qy 1995 GTGTCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCAATAACCCCTGAAGAATAAA 2054

Db 1 GGTTCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCAATAACCCCTGAAGAATAAA 60

Qy 2055 ACGGAATCAAGTCTTCTAGCTAAATTAGAGAAACTAAAGAGTATCAAGAACCCAGAGGTT 2114

Db 61 ACGGAATCAAGTCTTCTAGCTAAATTAGAGAAACTAAAGAGTATCAAGAACCCAGAGGTT 120

Qy 2115 CCAGAGAGTAACCAAGAACAGTGGCAAGCTAAGAGAAAAGTCAGAGTGTATTAAACCAGAAT 2174

Db 121 CCAGAGAGTAACCAAGAACAGTGGCAATCTAAGAGAAAAGTCAGAGTGTATTAAACCAGAAT 180

Qy 2175 CCTGCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAAAAGTTAATACAGAG 2234

Db 181 CCTGCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAAAAGTTAATACAGAG 240

Qy 2235 CAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAACACAGTCACCACCAATA 2294

Db 241 CAGAAACATACCACCTTTTGAGCAACCTGTCTTTTCAGTTTCAAAACACAGTCACCACCAATA 300

Qy 2295 TCAACATCTAAATGGTTTGACCCCAAAATCTATTGTAAAGACACCAAGCAGCAATACCTTG 2354

Db 301 TCAACATCTAAATGGTTTGACCCCAAAATCTATTGTAAAGACACCAAGCAGCAATACCTTG 360

Qy 2355 GATGATTACATGAGCTGTTTGAACCTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGT 2414

Db 361 GATGATTACATGAGCTGTTTGAACCTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGT 420

Qy 2415 CAGTTGTCAACACCTTTATGGCCAAACCTGCCTGTTTCCAGCAGCAACACGATCAAAATACTT 2474

Db 421 CAGTTGTCAACACCTTTATGGCCAAACCTGCCTGTTTCCAGCAGCAACACGATCAAAATACTT 480

Qy 2475 GCCACTCCACTTCAAAATTTACAGGTTTTCAGCATCTTCTTCCAGCAAAATGAATGCATTTTCG 2534

Db 481 GCCACTCCACTTCAAAATTTACAGGTTTTCAGCATCTTCTTCCAGCAAAATGAATGCATTTTCG 540

Qy 2535 GTTAAAGGAAGAATTTTATTCATATTTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGGTA 2594

Db 541 GTTANAGGAAGAATTTTATTCATATTTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGGTA 600

Qy 2595 TTTCAAGGTGTTAAATGAAGAAACAGATATATGCTATATAAAATATGTGAACCTTAGAAGAA 2654

Db 601 TTTCAAGGTGTTAAATGAAGAAACAGGATATATGCTATATAAAATATGTGAACCTTAGAAGAA 660

Qy 2655 GCAGATAACCAAACTCTTGATAGTTA-CCGGAACGAAATAGCTTTATTGAATAAACTACA 2713

Db 661 GCAGATAACCAAACTCTTGATAGTTA-CCGGAACGAAATAGCTTTATTGAATAAACTACA 720

Qy 2714 ACAACACAGT-GATAAGATCAT-CCGACTTTTATCATTTATGAAATCAGGACAGTACAT 2770

Db 721 ACAACNCAAGTGGATAAGATCATCCGACTTTTATGATTATGAAATCAGGACAGTACCT 779

RESULT 14

BG718184

LOCUS

DEFINITION BG718184 778 bp mRNA linear EST 08-MAY-2001 602696148F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4828282 5', mRNA sequence.

ACCESSION BG718184

VERSION BG718184.1 GI:13997371

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 778)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Plate: LLAM10745 row: p column: 11

FEATURES	source	High quality sequence stop: 770.
Location/Qualifiers		
1. .778		
/organism="Homo sapiens"		
/mol_type="mRNA"		
/db_xref="taxon:9606"		
/clone="IMAGE:4828282"		
/lab_host="DH10B"		
/clone_lib="NIH_MGC_97"		
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."		
ORIGIN		
Query Match	18.7%;	Score 723.4; DB 12; Length 778;
Best Local Similarity	98.8%;	Pred. No. 1.3e-142;
Matches 760;	Conservative 0;	Mismatches 6; Indels 3; Gaps 3;
QY	931	AGTTGGGTTCCATCTTTTCATTTCCCCAGTGCAGTTTTCTGTAGAAATGGAATCCGAGG 990
Db	5	
QY	991	ATTTAAGTGGCAGAGAAATTGACAATTGATTCATATGAACAAAGTGAGAGACATTAAAA 1050
Db	65	
QY	1051	ATAAGTTTAAAAATGAAGACCTTACTGATGAACTAAGCTTGAATAAAATTTCTGCTGATA 1110
Db	125	ATAAGTTTAAAAATGAAGACCTTACTGATGAACTAAGCTTGAATAAAATTTCTGCTGATA 184
QY	1111	CTACAGATAACTCGGGAAGTGTAAACCAAAATTATGATGATGGCAACCAACCCAGAGGACT 1170
Db	185	CTACAGATAACTCGGGAAGTGTAAACCAAAATTATGATGATGGCAACCAACCCAGAGGACT 244
QY	1171	GGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCGGCTAAGTGATGCTCTTTTAA 1230
Db	245	GGTTGAGTTTGTGCTCAAACTAGAGAAAAACAGTGTTCGGCTAAGTGATGCTCTTTTAA 304
QY	1231	ATAAATTGATTGGTCGTTACAGTCAAGCAATTGAAGCGCTTCCCCCAGATAAATATGGCC 1290
Db	305	ATAAATTGATTGGTCGTTACAGTCAAGCAATTGAAGCGCTTCCCCCAGATAAATATGGCC 364
QY	1291	AAATGAGAGTTTTTGTGCTAGAATTCAAGTGAGATTTGCTGAATTAAAGCTATTTCAAGAGC 1350
Db	365	AAATGAGAGTTTTTGTGCTAGAATTCAAGTGAGATTTGCTGAATTAAAGCTATTTCAAGAGC 424
QY	1351	CAGATGATGCACGTGACTACTTTCAAATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTG 1410
Db	425	CAGATGATGCACGTGACTACTTTCAAATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTG 484
QY	1411	TTCATATATCTTTTGCACAAAATTGGAACCTGTCAAGTAATGTCAAA - AAAAGTAAACAA 1469
Db	485	TTCATATATCTTTTGCACAAAATTGGAACCTGTCAAGTAATGTCAAA - AAAAGTAAACAA 544
QY	1470	CTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCACTAGAAATGCTGGAAATTTGCCCTG 1529
Db	545	CTTCTTCAAAAAGCTGTAGAACGTGGAGCAGTACCACTAGAAATGCTGGAAATTTG - CCTG 603
QY	1530	CGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTTTCAGAGGAGGAA - AAGAAGAAATTTATC 1588
Db	604	CGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTTTCAGAGGAGGAA - AAGAAGAAATTTATC 663
QY	1589	AGCATCTACGGTATTAACTGCCCAAGAATCATTTTCCGGTTCACTTGGGCAATTTACAGAA 1648
Db	664	AGCATCTACGGTATTAACTGGCCCAAGAATCATTTTCCGGTTCACTTGGGCAATTTACAGAA 723

QY 1649 TAGGAACAACAGTGTGATTCCAGAGGACAGACTACTAAAGCCAGGTTT 1697
||||| |

Db 724 TAGGAACAAAAGCTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT 772
||||| |

RESULT 15
AU132533 linear EST 01-AUG-2002

LOCUS AU132533 NT2RP4 Homo sapiens cDNA clone NT2RP4000021 5', mRNA

DEFINITION sequence.

VERSION AU132533.1 GI:10992887

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 821)
Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuho,Y. and
Isogai,T.

TITLE HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T.,
Sugano,S., Masuho,Y., Isogai,T.)

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1..821
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP4000021"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RP4"
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Query Match 18.6%; Score 720.6; DB 9; Length 821;
Best Local Similarity 95.1%; Pred. No. 5.2e-142;
Matches 774; Conservative 0; Mismatches 37; Indels 3; Gaps 3;

QY 897 GTGAATGTAAGACTGTACAGACTTCCTAGAAAAACAGTTTGCGTTCCATCTTTTCATTTC 956
||| |

Db 5 GGCACGGGAAAATTCAAACGTGTTTGGCGAAAGGAGTTTGCGTTCCATCTTTTCATTTC 64
||| |

QY 957 CCAGTGCAGTTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAAATT 1016
||| |

Db 65 CCAGCGCAGCTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAAATT 124
||| |

QY 1017 GATTCATAATGAACAAAAGTGAGAGACATTAATAAATAAGTTTAAAAATGAAGACCTTACT 1076
||| |

Db 125 GATTCATAATGAACAAAAGTGAGAGACATTAATAAATAAGTTTAAAAATGAAGACCTTACT 184
||| |

QY 1077 GATGAACATAAGCTTGAATAAAAAATTTCTGCTGATACACAGATAACTCGGGAAGTTGTTAAC 1136
||| |

Db 185 GATGAACATAAGCTTGAATAAAAAATTTCTGCTGATACACAGATAACTCGGGAAGTTGTTAAC 244
||| |

QY 1137 CAAATTATGATGATGGCAAAACACCAGAGGACTGGTTGAGTTTGTTGCTCAAACTAGAG 1196
||| |

Db 245 CAAATTATGATGATGGCAAAACACCAGAGGACTGGTTGAGTTTGTTGCTCAAACTAGAG 304
||| |

QY	1197	AAAAACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATGAATGGTCGTTACAGTCAA	1256
Db	305	AAAAACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATGAATGGTCGTTACAGTCAA	364
QY	1257	GCAATTGAAGCGCTTCCCCCAGATAAATATGGCCAAAATGAGAGTTTTGCTAGAAATCAA	1316
Db	365	GCAATTGAAGCGCTTCCCCCAGATAAATATGGCCAAAATGAGAGTTTTGCTAGAAATCAA	424
QY	1317	GTGAGATTTGCTGAATTAAAAAGCTATTCAAGAGCCAGATGACGCTGACTACTTTCAA	1376
Db	425	GTGAGATTTGCTGAATTAAAAAGCTATTCAAGAGCCAGATGACGCTGACTACTTTCAA	484
QY	1377	ATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTGCATATATCTTTTGACAAATTTGAA	1436
Db	485	ATGGCCAGAGCAAACTGCAAGAAATTTGCTTTTGCATATATCTTTTGACAAATTTGAA	544
QY	1437	CTGTCAACAAGGTAATGTCAAAAAAAGTAAACAACCTTCTTCAAAAAAGCTGTAGAACGTGGA	1496
Db	545	CTGTCAACAAGGTAATGTCAAAAAAAGTAAACAACCTTCTTCAAAAAAGCTGTAGAACGTGGA	604
QY	1497	GCAGTACCACTAGAAATGCTGGAAATTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAG	1556
Db	605	GCAGTACCACTAGAAATGCTGGAAATTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAG	664
QY	1557	CTGCTTTTCAGAGGAGGAAAAGAGAATTTATCAGCATCTACGGTATTAACTGCCCAAGAA	1616
Db	665	CTGCTTTTCAGAGGAGGAAAAGAGAATTTATCAGCATCTACGGTATTAACTGCCCAAGAA	724
QY	1617	TCATTTTCCGGTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCACAGAGGA	1676
Db	725	TCATTTTCCGG-TCACTTGGGCATTT-CAGAAATAGG-ACCACAGTTGTGATTNCAGANGA	781
QY	1677	CAGACTACTAAAGCCAGGTTTTTTATATGGAGAGA	1710
Db	782	CAGACTCTAAAGNCAGGTTTTTATATGGAGAGACA	815

This Page Blank (uspto)

Db	1321	GATTTGCTGAATTA	AAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCA	ATGG	1380
QY	1381	CCAGAGCAAAC	TGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGA	ACTGT	1440
Dd	1381	CCAGAGCAAAC	TGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGA	ACTGT	1440
QY	1441	CACAAAGGTAATGT	CAAAAAAAGTAAACAACCTCTTCAAAAAAGCTGTAGAACGTGGAGCAG	1500	
Dd	1441	CACAAAGGTAATGT	CAAAAAAAGTAAACAACCTCTTCAAAAAAGCTGTAGAACGTGGAGCAG	1500	
QY	1501	TACCACTAGAAATGCT	TGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC	1560	
Dd	1501	TACCACTAGAAATGCT	TGGAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC	1560	
QY	1561	TTTCAGAGGAGGAAAAA	AGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCAT	1620	
Dd	1561	TTTCAGAGGAGGAAAAA	AGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCAT	1620	
QY	1621	TTTCCGGTTCACT	TGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA	1680	
Dd	1621	TTTCCGGTTCACT	TGGGCATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA	1680	
QY	1681	CTACTAAAGCCAG	TTTTATATGGAGAGAACATGCCACCAAGATGCAGAAATAGGTT	1740	
Dd	1681	CTACTAAAGCCAG	TTTTATATGGAGAGAACATGCCACCAAGATGCAGAAATAGGTT	1740	
QY	1741	ACCGGAATTCATT	TGAGACAAAACCTAACAAAACTAAAACAGTCATGCCCATTTGGAAGAGTCC	1800	
Dd	1741	ACCGGAATTCATT	TGAGACAAAACCTAACAAAACTAAAACAGTCATGCCCATTTGGAAGAGTCC	1800	
QY	1801	CAGTTAACCTTCT	AAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT	1860	
Dd	1801	CAGTTAACCTTCT	AAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT	1860	
QY	1861	GTTTTATGAAAAGAC	AAAACCTCTAGATCAGAAATGCCAGATTTGGTTGTGCCTGGATCTA	1920	
Dd	1861	GTTTTATGAAAAGAC	AAAACCTCTAGATCAGAAATGCCAGATTTGGTTGTGCCTGGATCTA	1920	
QY	1921	AACCAAGTGGAAT	GTANTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT	1980	
Dd	1921	AACCAAGTGGAAT	GTANTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCATT	1980	
QY	1981	TCAAGGAACCTCT	GTGTGATGAAAAGAGTTCTGAACTTATTACTGATTCAAATAA	2040	
Dd	1981	TCAAGGAACCTCT	GTGTGATGAAAAGAGTTCTGAACTTATTACTGATTCAAATAA	2040	
QY	2041	CCCTGAAGAATAAA	ACCGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC	2100	
Dd	2041	CCCTGAAGAATAAA	ACCGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC	2100	
QY	2101	AAGAACCCAGAG	GTCCAGAGAGTAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT	2160	
Dd	2101	AAGAACCCAGAG	GTCCAGAGAGTAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT	2160	
QY	2161	GTATTAAACCAGA	ATCCTGCTGCATCTTCAAATCAGTGGCAGATTCGGAGTTAGCCCGAA	2220	
Dd	2161	GTATTAAACCAGA	ATCCTGCTGCATCTTCAAATCAGTGGCAGATTCGGAGTTAGCCCGAA	2220	
QY	2221	AAGTTAATACAG	AGCAGAAAACATACCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAC	2280	
Dd	2221	AAGTTAATACAG	AGCAGAAAACATACCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAC	2280	
QY	2281	AGTCACCACCA	ATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA	2340	
Dd	2281	AGTCACCACCA	ATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA	2340	
QY	2341	GCAGCAATACCT	TGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAATGACT	2400	
Dd	2341	GCAGCAATACCT	TGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAATGACT	2400	
QY	2401	TTCCACCTGCT	TGTCAGTTGTCAACACCTTATGGCCAACTGCCTGTTTTCCAGCAGCAAC	2460	

QY 3541 GAAAAAATGATTGTCAGTTATTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600
Db |||||||
QY 3541 GAAAAAATGATTGTCAGTTATTCGTAATGTCAGATAGGAGGTATAAAATATATTGGACT 3600
Db |||||||
QY 3601 GTTATACTCTTGAATCCCTGTGGAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db |||||||
QY 3601 GTTATACTCTTGAATCCCTGTGGAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db |||||||
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTTAAAGAAAACTGTAAATAGCAACCAC 3720
Db |||||||
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTTAAAGAAAACTGTAAATAGCAACCAC 3720
Db |||||||
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTCTCTCTCTTTATGCTCTTGTGTAATCTAC 3780
Db |||||||
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTCTCTCTCTTTATGCTCTTGTGTAATCTAC 3780
Db |||||||
QY 3781 TTGACATCATTTTACTCTTTGGAATAGTGGGTGGATAGCAAGTATATTCTTAAAAAATTTG 3840
Db |||||||
QY 3781 TTGACATCATTTTACTCTTTGGAATAGTGGGTGGATAGCAAGTATATTCTTAAAAAATTTG 3840
Db |||||||
QY 3841 TAAATAAAGTTTGTGGCTAAAATGA 3866
Db |||||||
QY 3841 TAAATAAAGTTTGTGGCTAAAATGA 3866
Db |||||||

RESULT 2

ABV73974
ID ABV73974 standard; cDNA; 3866 BP.
XX
AC ABV73974;
XX
DT 13-JAN-2003 (first entry)
XX
DE Human tyrosine threonine kinase cDNA.
XX
KW Tyrosine threonine kinase; TTK; protein kinase; enzyme; human;
KW mitotic checkpoint; colon cancer; breast cancer; tumour; cytostatic;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
FH Key
FT CDS 1026..3551
FT /tag= a
FT /product= "TTK"
XX

PN WO200268444-A1.
XX
XX
PD 06-SEP-2002.
XX
PF 21-FEB-2002; 2002WO-US005278.
XX
PR 21-FEB-2001; 2001US-0271254P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Reinhard C, Jefferson AB, Chan VW;
XX
DR WPI; 2002-698650/75.
DR P-PSDB; ABP54938.
XX

PT Reducing growth of cancer cells comprises reducing Tyrosine Threonine
PT Kinase (TTK) activity, useful in diagnosing and treating disorders with
PT abnormal expression levels and activity of TTK, such as lung, colon,
PT prostate and ovarian cancer.
XX
PS Disclosure; Page 67-71; 113pp; English.
XX
CC The present sequence is that of cDNA encoding human tyrosine threonine
CC kinase (TTK), a mixed specificity protein kinase that is involved in the
CC cell cycle spindle assembly checkpoint. The invention is based on the
CC finding that TTK is differentially expressed in colon tumour cells

CC relative to healthy tumour colon cells as detected by microarray
CC analysis. Differential expression was confirmed in cell lines derived
CC from various forms of cancer, indicating involvement of TTK in cancer as a
CC more general mechanism. Disruption of TTK function using antisense
CC oligonucleotides decreased proliferation, inhibited anchorage-independent
CC growth, and induced apoptosis of cancer cell lines, including a
CC metastatic breast cancer cell line (MDA-MB-213) and a colorectal
CC carcinoma cell line (SW620). These findings serve as the basis for the
CC materials and methods of the invention of appropriate therapeutic and/or
CC patient, for planning and selection of appropriate treatment of cancer associated with
CC prophylactic treatment, and for treatment of cancer associated with
CC aberrant TTK levels, e.g. associated with overexpression or
CC overproduction of TTK, by inhibition of gene product production (by
CC decreasing levels of transcription and/or translation), by decreasing TTK
CC activity by decreasing TTK gene product production, and/or by reducing
CC one or more of TTK's kinase activities in cancer cells, especially breast
CC cancer or colon cancer cells. The present invention has the advantage of
CC having the ability to project disease progression based on expression of
CC TTK in a malignant or pre-malignant growth. It allows a more systematic
CC approach for intervention of a cancerous disease

XX
SQ Sequence 3866 BP; 1268 A; 731 C; 752 G; 1115 T; 0 U; 0 Other;
Query Match 100.0%; Score 3866; DB 6; Length 3866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATTCCTTTTCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTG 60
Db |||||||
QY 61 CAATGGCACAATCTCAGCTTACTGCAACCTCCGCTCCCGGTTTCAAGCGATTTCTCTGTC 120
Db |||||||
QY 121 CTCAGCCTCTCAAGTAGCTGGGATTACAGGCAATGTCACCCCTGGCTAACTAATT 180
Db |||||||
QY 181 CTTTCTATTTAGTAGAGATGGGTTTTCACCATGTTGTTGAGGCTGTTTGAACCTCTG 240
Db |||||||
QY 241 ACCTCAGGTGATCCACTTGGCTTGGCTCCCAAGAGTCTAGGATTACAGCCGTGAAACTG 300
Db |||||||
QY 301 TGCCTGGCTGATCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 360
Db |||||||
QY 361 AGGCTGGAGTGCAGTGGTGGATCTTGGCTCACTATAACCTCCACCTCTGTTTCAAGT 420
Db |||||||
QY 421 GATCTCCCACTTTAGCTCTCTGAGTAGCTGTGATTACAGGCGTGACACACACCCGG 480
Db |||||||
QY 481 CTAATTTTGTATTTTATTAGAGACAGGGTTTCAACCATGTTGGCCAGGCTGTTCTCAA 540
Db |||||||
QY 541 CTCCTGGACTCAAGGATCCGCTGCTTCCACTCCCAAGTCCCGAGATTACAGGTGTG 600
Db |||||||
QY 601 AGTCACCATGCTGACCTTATAATTTCTTAAGTCATTTTCTGTTGTTCTTCTTAG 660
Db |||||||
QY 661 GGTCCTCACAACAAATCTGCATTAGGCGGTACAATAATCTTAACTCATGATTCACAAA 720
Db |||||||

Db . 661 GGCTCTCACAACAAATCTGCATTAGCGGTACAAATAATCCTTAACTTCATGATTCACAAA 720
Qy 721 AGGAAGATGAAGTGATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT 780
Db 721 AGGAAGATGAAGTGATTCATGATTTAGAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT 780
Qy 781 GGATGATGATCCTTAATCCAGATACAGTAAATAATGGGGTATGCGAAGGTAGAAATACAAA 840
Db 781 GGATGATGATCCTTAATCCAGATACAGTAAATAATGGGGTATGCGAAGGTAGAAATACAAA 840
Qy 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAAAACATTTTGGATACATTTGTTGATGTA 900
Db 841 TTTGGTTTAAATTAATTAATCTAAATATCTAAAAACATTTTGGATACATTTGTTGATGTA 900
Qy 901 ATGTAAGACTGTACAGACTTCCTAGAAAAACAGTTTGGGTTCCTATCTTTTCATTTCCCCAG 960
Db 901 ATGTAAGACTGTACAGACTTCCTAGAAAAACAGTTTGGGTTCCTATCTTTTCATTTCCCCAG 960
Qy 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAATTGATT 1020
Db 961 TGCAGTTTCTGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAATTGATT 1020
Qy 1021 CCATAATGAACAAAGTGAGAGACATTTAAAAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
Db 1021 CCATAATGAACAAAGTGAGAGACATTTAAAAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
Qy 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACTACAGATAAATCGGAACTGTTAAACCAA 1140
Db 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACTACAGATAAATCGGAACTGTTAAACCAA 1140
Qy 1141 TTATGATGATGGCAAAACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAA 1200
Db 1141 TTATGATGATGGCAAAACAAACCCAGAGGACTGGTTGAGTTTGTGCTCAAACTAGAGAAAA 1200
Qy 1201 ACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATGATTGTTGCTCAAACTAGAGAAAA 1260
Db 1201 ACAGTGTCCGCTAAGTGATGCTCTTTTAAATAAATGATTGTTGCTCAAACTAGAGAAAA 1260
Qy 1261 TTGAAGCGCTTCCCCAGATAAAATATGGCCAAAATGAGAGTTTGTGCTCAAACTAGAGAAAA 1320
Db 1261 TTGAAGCGCTTCCCCAGATAAAATATGGCCAAAATGAGAGTTTGTGCTCAAACTAGAGAAAA 1320
Qy 1321 GATTTGCTGAATTAAGCTATTTCAAGAGCCAGATGATGCACGTGACTACTTTCAAAATGG 1380
Db 1321 GATTTGCTGAATTAAGCTATTTCAAGAGCCAGATGATGCACGTGACTACTTTCAAAATGG 1380
Qy 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTGAACTGT 1440
Db 1381 CCAGAGCAAACTGCAAGAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTGAACTGT 1440
Qy 1441 CACAAGGTAATGTCAAAAAAAGTAAACAACCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAATGTCAAAAAAAGTAAACAACCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Qy 1501 TACCAC TAGAAATGCTGGAAATTTGCCCTGCGGAAATTTAAACCTCAAAAAAAGCAGCTGC 1560
Db 1501 TACCAC TAGAAATGCTGGAAATTTGCCCTGCGGAAATTTAAACCTCAAAAAAAGCAGCTGC 1560
Qy 1561 TTTCAGAGGAGGAAAGAAAGAAATTTATCAGCATCTACGGTATTAACCTGCCCCAAGAAATCAT 1620
Db 1561 TTTCAGAGGAGGAAAGAAAGAAATTTATCAGCATCTACGGTATTAACCTGCCCCAAGAAATCAT 1620
Qy 1621 TTTCCGGTTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTTGATTCAGAGGACAGA 1680
Db 1621 TTTCCGGTTTCACTTGGGCATTTACAGAAATAGGAACAACAGTTTGATTCAGAGGACAGA 1680
Qy 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAATGCCACCACAGATGCAGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAAATGCCACCACAGATGCAGAAATAGGTT 1740
Qy 1741 ACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCATTTGGAAGAGTCC 1800
Db 1741 ACCGGAATTCATTGAGACAAACTAAACAAACTAAACAGTCATGCCCATTTGGAAGAGTCC 1800

Qy 1801 CAGTTAAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAAACCTTCTAAATAGCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Qy 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCAGATTTGGTTGTCCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCAGATTTGGTTGTCCTGGATCTA 1920
Qy 1921 AACCAAGTGGAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAAATAGTCATT 1980
Qy 1981 TCAAGGAAACCTCTGGTGTGAGATGAAAAGAGTTCTGAACTTATTATTACTGATTCAATAA 2040
Db 1981 TCAAGGAAACCTCTGGTGTGAGATGAAAAGAGTTCTGAACTTATTATTACTGATTCAATAA 2040
Qy 2041 CCCTGAAGAATAAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
Db 2041 CCCTGAAGAATAAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC 2100
Qy 2101 AAGAACCAAGAGGTTCCAGAGAGTAACCCAGAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGT 2160
Db 2101 AAGAACCAAGAGGTTCCAGAGAGTAACCCAGAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGT 2160
Qy 2161 GTATTAAACCAAGATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACCAAGATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGTTAGCCCGAA 2220
Qy 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAC 2280
Qy 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCCAAAATCTATTGTGAAGACACCAA 2340
Db 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCCAAAATCTATTGTGAAGACACCAA 2340
Qy 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAGAACTCCAGTTGTAAAGAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTAGAACTCCAGTTGTAAAGAATGACT 2400
Qy 2401 TTCCACCTGCTTGTGATGTCAACACCTTATGGCCAACTGCTGCTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGATGTCAACACCTTATGGCCAACTGCTGCTTTCCAGCAGCAAC 2460
Qy 2461 AGCATCAAATACTTGCCACTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAATACTTGCCACTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
Qy 2521 ATGAATGCAATTCGGTTTAAAGGAAGAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCAATTCGGTTTAAAGGAAGAATTTATTCATATTTAAAGCAGATAGGAAGTGGAG 2580
Qy 2581 GTTCAAGCAAGGTATTTCAAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTCAAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAAATATG 2640
Qy 2641 TGAACCTTAGAAGAACAGATAAACCAACTCTTGTAGTAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAACAGATAAACCAACTCTTGTAGTAGTTACCGGAACGAAATAGCTTATT 2700
Qy 2701 TGAATAAACTACAACAACAGATGATAAGATCATCCGACTTTATGATATGAAATCACGG 2760
Db 2701 TGAATAAACTACAACAACAGATGATAAGATCATCCGACTTTATGATATGAAATCACGG 2760
Qy 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTTAATAGTTGGCTTAAAA 2820
Qy 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880

QY 2101 AAGAACCCAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db |||||
QY 2101 AAGAACCCAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db |||||
QY 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Db |||||
QY 2161 GTATTAAACAGAAATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGGAGTTAGCCCGAA 2220
Db |||||
QY 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGGCAACCTGTCTTTTTCAGTTTCAAAAC 2280
Db |||||
QY 2221 AAGTTAATACAGAGCAGAAACATACCACTTTTGGCAACCTGTCTTTTTCAGTTTCAAAAC 2280
Db |||||
QY 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGGACCCAAATCTATTGTGAAGACACCAA 2340
Db |||||
QY 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGGACCCAAATCTATTGTGAAGACACCAA 2340
Db |||||
QY 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTGTAGAACTCCAGTTGTAAAGAAATGACT 2400
Db |||||
QY 2401 TTCCACCTGCTGTGTCAGTTGTCAACACCTTATGGCCAACTGCTGCTGTTTCCAGCAGCAAC 2460
Db |||||
QY 2461 AGCATCAATACTTGGCACTCCACTTCAAAATTTACAGGTTTACAGTCTTCTTTCAGCAA 2520
Db |||||
QY 2521 ATGAATGCAATTCGGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
Db |||||
QY 2581 GTTCAAGCAAGGTATTTCAAGGTGTAAATGAAAGAAACAGATATATGCTATAAAATATG 2640
Db |||||
QY 2641 TGAAGCAAGGTATTTCAAGGTGTAAATGAAAGAAACAGATATATGCTATAAAATATG 2700
Db |||||
QY 2701 TGAATAAATCAACACACAGTGATAAGATCATCCGACTTTATGATTGAAATCACGG 2760
Db |||||
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAA 2820
Db |||||
QY 2821 AGAAAAATCCATTGATCCATGGGAACCGAAGAGTTACTGGAAAAATATTGATCTTAATAGTTGGCTTAAA 2880
Db |||||
QY 2881 TTCACACAATCCATCAACATGGCATGTTTCACAGTGATCTTAAACCGACTAATCTTCTGA 2940
Db |||||
QY 2941 TAGTTGATGAATGCTAAAGCTAAATGATTTTGGGATTCGAAACCAATGCAACCCAGATA 3000
Db |||||
QY 3001 CAACAAGTGTGTAAAGATTCTCAGGTTGGCAGAGTTAATATATGCCACCAGAAGCAA 3060
Db |||||
QY 3061 TCAAGATATGTCTTCTCCAGAGAGAATGGGAATCTAAGTCAAAAGTAAGCCCCAAAA 3120
Db |||||
QY 3121 GTGATGTTTGGTCTTAGGATGTAATTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db |||||

QY 3181 AGCAGATAATTAATCAGATTTTCTAAATACATGCCATAATTGATCCTAATCATGAAATTG 3240
Db |||||
QY 3241 AATTTCCCGATATTTCCAGAGAAAGATCTTCAAGATGTGTAAAGTGTGTTTAAAGGG 3300
Db |||||
QY 3301 ACCCAAAACAGAGGATATCCATTCTGAGCTCCTGGCTCATCCATATGTCAAATTCAAA 3360
Db |||||
QY 3361 CTCATCCAGTTAAACCAATGGCCAAAGGGAACCACTGAAAGAAATGAAATATGTTCTGGGCC 3420
Db |||||
QY 3421 AACTTGTGGTCTGAATTTCTCTAACTCCATTCTGAAAGCTGCTAAACTTTATATGAAC 3480
Db |||||
QY 3481 ACTATAGTGGTGTGAAAGTCAATAATCTTCTCACTCCCAAGACTTTTGAAAAAAGGG 3540
Db |||||
QY 3541 GAAAAAATGATTTGCAGTTATTCGTAATGTGAGATAGGAGGTATAAAATATATTGGACT 3600
Db |||||
QY 3601 GTTATACTCTTGAATCCCTGTGGAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db |||||
QY 3661 ATCAGCAAAAAAATTTCACTGAGATTATCTTTTAAAGAAAACTGTAAAAATAGCAACCAC 3720
Db |||||
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTCTCTGTTTCTGTTTATGCTCTGTGTAATCTAC 3780
Db |||||
QY 3781 TTGACATCATTTTACTCTTTGGAATAGTGGGTGGATAGCAAGTATATTCTAAAAAATTG 3840
Db |||||
QY 3841 TAAATAAAGTTTGTGGCTAAAATGA 3866
Db |||||

RESULT 4
ADD49931
ID ADD49931 standard; cDNA; 3866 BP.
XX
AC ADD49931;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human lung specific tumour antigen L 1479p cDNA.
XX
KW Human; ss; lung cancer antigen; cytostatic; lung cancer; gene therapy;
KW vaccine; T-cell; gene.
XX
OS Homo sapiens.
XX
PN US2003194764-A1.
XX
PD 16-OCT-2003.
XX
PF 04-APR-2002; 2002US-00116712.
XX
PR 05-APR-2001; 2001US-0282289P.
XX
PR 05-OCT-2001; 2001US-0327511P.
XX

QY 1261 TTGAAGCGCTTCCCCAGATAAATATGGCCAAAATGACAGTTTTGCTAGATTCAAGTGA 1320
Db 1261 TTGAAGCGCTTCCCCAGATAAATATGGCCAAAATGACAGTTTTGCTAGATTCAAGTGA 1320
QY 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGACGCTGACTACTTTCAAAATGG 1380
Db 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGACGCTGACTACTTTCAAAATGG 1380
QY 1381 CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGGCACAATTTGAACTGT 1440
Db 1381 CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGGCACAATTTGAACTGT 1440
QY 1441 CACAAGGTAATGTCAAAAAAAGTAACAACCTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAATGTCAAAAAAAGTAACAACCTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
QY 1501 TACCACTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
Db 1501 TACCACTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
QY 1561 TTTTCAGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCAT 1620
Db 1561 TTTTCAGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAATCAT 1620
QY 1621 TTTCCGGTTCACTTGGGCATTTACAGAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680
Db 1621 TTTCCGGTTCACTTGGGCATTTACAGAATAGGAACAACAGTTGTGATTCAGAGGACAGA 1680
QY 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAACATGCCACCACAAGATGCGAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTATATGGAGAGAACATGCCACCACAAGATGCGAAATAGGTT 1740
QY 1741 ACCGGAATTCATTGAGACAAAACCTAACAAAACCTAACAGTCATGCCCAITTTGGAAGAGTCC 1800
Db 1741 ACCGGAATTCATTGAGACAAAACCTAACAAAACCTAACAGTCATGCCCAITTTGGAAGAGTCC 1800
QY 1801 CAGTTAACCTTCTAAATAGCCCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAACCTTCTAAATAGCCCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
QY 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCGCAGATTGGTTGTCCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCGCAGATTGGTTGTCCTGGATCTA 1920
QY 1921 AACCAAGTGAAATGATTCCTGTGAATTAAGAAAATTAAGAACTCTGTTCAAAAATAGTCATT 1980
Db 1921 AACCAAGTGAAATGATTCCTGTGAATTAAGAAAATTAAGAACTCTGTTCAAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGGTGTCAAGATGAAAAGAGTTCTGAACCTTATTAATCTGATTCATAA 2040
Db 1981 TCAAGGAACCTCTGGTGTCAAGATGAAAAGAGTTCTGAACCTTATTAATCTGATTCATAA 2040
QY 2041 CCTGGAAGATAAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAAGAGTATC 2100
Db 2041 CCTGGAAGATAAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAAGAGTATC 2100
QY 2101 AAGAACCCAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACCCAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
QY 2161 GTATTAAACAGAAATCCTGTGCACTCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACAGAAATCCTGTGCACTCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGGCAACCTGTCTTTTCAGTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAACATACCACCTTTTGGCAACCTGTCTTTTCAGTTTCAAAAC 2280
QY 2281 AGTCACCACCAATATCAACATCTAAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA 2340
Db 2281 AGTCACCACCAATATCAACATCTAAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAAATGACT 2400

Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAAATGACT 2400
QY 2401 TTCCACCTGCTTGTGTCAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGTCAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAATACTTGCCTTAAAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2520
Db 2461 AGCATCAATACTTGCCTTAAAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2520
QY 2521 ATGAATGCATTTGCTTAAAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCATTTGCTTAAAGGAAGAATTTATTCATATTAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
QY 2641 TGAACCTTAGAAGAAGCAGATTAACCAAACTCTTGTAGATTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAAGCAGATTAACCAAACTCTTGTAGATTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAACAACACAGTGAAGATCATCCGACTTTTATGATTATGAAATCACGG 2760
Db 2701 TGAATAAACTACAACAACACAGTGAAGATCATCCGACTTTTATGATTATGAAATCACGG 2760
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAAA 2820
QY 2821 AGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGCAG 2880
QY 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCCAGCTAACTTTCTGA 2940
Db 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCCAGCTAACTTTCTGA 2940
QY 2941 TAGTTGATGGAATGCTAAAGTAATTTGATTTGGGATGCAAAACCAATGCAACCAAGATA 3000
Db 2941 TAGTTGATGGAATGCTAAAGTAATTTGATTTGGGATGCAAAACCAATGCAACCAAGATA 3000
QY 3001 CAACAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATATATGCCCACAGAAAGCAA 3060
Db 3001 CAACAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATATATGCCCACAGAAAGCAA 3060
QY 3061 TCAAAGATATGCTTCTCCAGAGAGAAATGGGAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
Db 3061 TCAAAGATATGCTTCTCCAGAGAGAAATGGGAATCTAAGTCAAAAGATAAGCCCCAAAA 3120
QY 3121 GTGATGTTTGGTCTTAGGATGATTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCTTAGGATGATTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
QY 3181 AGCAGATAATTAATCAGATTCTAAATTTACATGCCATAATTGATCCTAATCATGAAATTTG 3240
Db 3181 AGCAGATAATTAATCAGATTCTAAATTTACATGCCATAATTGATCCTAATCATGAAATTTG 3240
QY 3241 AATTTCCGATATTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCGATATTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
QY 3301 ACCCAAAACAGAGGATATCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db 3301 ACCCAAAACAGAGGATATCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
QY 3361 CTCATCCAGTTAACCAATGGCCAAAGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAACCAATGGCCAAAGGAACCACTGAAGAAATGAATATGTTCTGGGCC 3420
QY 3421 AACTTGTGGTCTGAATTTCTCCTAACTCCATTGAAAGCTGCTAAACTTTTATATGAAC 3480

Db 541 CTCTGGACTCAAGGGATCCGCTGCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTG 600
QY 601 AGTCACCATGCTGACCTTATAATTCTTAAGTCAATTTTCTCTGGTCCATTTCTCCTTAG 660
Db 601 AGTCACCATGCTGACCTTATAATTCTTAAGTCAATTTTCTCTGGTCCATTTCTCCTTAG 660
QY 661 GGTCTCTACAAACAAATCTGCATTAGGCGGTACAATAATCCTTAACCTTCATGATTCACAAA 720
Db 661 GGTCTCTACAAACAAATCTGCATTAGGCGGTACAATAATCCTTAACCTTCATGATTCACAAA 720
QY 721 AGGAAGATGAAGTGATTCATGATTTAGAAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT 780
Db 721 AGGAAGATGAAGTGATTCATGATTTAGAAAAGGGGAAGTAGTAAGCCCACTGCACACTCCT 780
QY 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAACAAA 840
Db 781 GGATGATGATCCTAAATCCAGATACAGTAAATAATGGGGTATGGGAAGGTAGATAACAAA 840
QY 841 TTTGGTTTAAATTAATTATCTAAATATCTAAAAACATTTTGGATACATTTGTGATGTA 900
Db 841 TTTGGTTTAAATTAATTATCTAAATATCTAAAAACATTTTGGATACATTTGTGATGTA 900
QY 901 ATGTAAGACTGTACAGACTTCTCTAGAAAAACAGTTTGGGTCCATCTTTTCATTTCCCCAG 960
Db 901 ATGTAAGACTGTACAGACTTCTCTAGAAAAACAGTTTGGGTCCATCTTTTCATTTCCCCAG 960
QY 961 TGCAGTTTCTGTAGAAAATGGAAATCCGAGGATTTAAGTGGCAGAGAATTCACAATTGATT 1020
Db 961 TGCAGTTTCTGTAGAAAATGGAAATCCGAGGATTTAAGTGGCAGAGAATTCACAATTGATT 1020
QY 1021 CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
Db 1021 CCATAATGAACAAAGTGAGAGACATTAATAATAAGTTTAAAAATGAAGACCTTACTGATG 1080
QY 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACTACAGATAACTCGGGAACCTGTTAAACCAA 1140
Db 1081 AACTAAGCTTGAATAAAATTTCTGCTGATACTACAGATAACTCGGGAACCTGTTAAACCAA 1140
QY 1141 TTATGATGATGGCAAAACAAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
Db 1141 TTATGATGATGGCAAAACAAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAA 1200
QY 1201 ACAGTGTTCGCTAAGTGATGCTCTTTTAAATAAATGATGGTTCGTTACAGTCAAGCAA 1260
Db 1201 ACAGTGTTCGCTAAGTGATGCTCTTTTAAATAAATGATGGTTCGTTACAGTCAAGCAA 1260
QY 1261 TTGAAGCGCTTCCCCAGATAAATATGGCCAAATGAGAGTTTGTCTAGAAATTCAGTGA 1320
Db 1261 TTGAAGCGCTTCCCCAGATAAATATGGCCAAATGAGAGTTTGTCTAGAAATTCAGTGA 1320
QY 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTTCAAAATGG 1380
Db 1321 GATTTGCTGAATTAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTTCAAAATGG 1380
QY 1381 CCAGAGCAAACTGCAAGAAAATTTGCTTTTGCATATATCTTTTGCACAAATTTGAACGTG 1440
Db 1381 CCAGAGCAAACTGCAAGAAAATTTGCTTTTGCATATATCTTTTGCACAAATTTGAACGTG 1440
QY 1441 CACAAGGTAATGTCAAAAAAAGTAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
Db 1441 CACAAGGTAATGTCAAAAAAAGTAACAACTTCTTCAAAAAGCTGTAGAACGTGGAGCAG 1500
QY 1501 TACCACCTAGAAAATGCTGGAAATGCCCCTGGGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
Db 1501 TACCACCTAGAAAATGCTGGAAATGCCCCTGGGGAATTTAAACCTCCAAAAAAGCAGCTGC 1560
QY 1561 TTTTCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACCTCCCAAGAAATCAT 1620
Db 1561 TTTTCAGAGGAGGAAAAGAAATTTATCAGCATCTACGGTATTAACCTCCCAAGAAATCAT 1620
QY 1621 TTTCCGGTTCACTTGGGCATTTTACAGAAATAGGAACAAACAGTTTGTGATTCAGAGGACAGA 1680
Db 1621 TTTCCGGTTCACTTGGGCATTTTACAGAAATAGGAACAAACAGTTTGTGATTCAGAGGACAGA 1680

QY 1681 CTACTAAAGCCAGGTTTTTTATATGGAGAGAAACATGCCACCACAAGATGCAGAAAATAGGTT 1740
Db 1681 CTACTAAAGCCAGGTTTTTTATATGGAGAGAAACATGCCACCACAAGATGCAGAAAATAGGTT 1740
QY 1741 ACCGGAATTCATTGAGACAAAACCTAACAAAACTAAACAGTCATGCCCATTTGGAAAGAGTCC 1800
Db 1741 ACCGGAATTCATTGAGACAAAACCTAACAAAACTAAACAGTCATGCCCATTTGGAAAGAGTCC 1800
QY 1801 CAGTTAACCTTCTAATAATAGCCCAAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
Db 1801 CAGTTAACCTTCTAATAATAGCCCAAGATTGTGATGTGAAGACAGATGATTCAGTTGTACCTT 1860
QY 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTA 1920
Db 1861 GTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTA 1920
QY 1921 AACCAAGTGGAAAATGATTCCTGTGAATTAAGAAAATTAAGAGTCTGTTCAAAAATAGTCATT 1980
Db 1921 AACCAAGTGGAAAATGATTCCTGTGAATTAAGAAAATTAAGAGTCTGTTCAAAAATAGTCATT 1980
QY 1981 TCAAGGAACCTCTGGTGTGAGATGAAAAGAGTTCTGAACTTATTATTACTGATTTCAATAA 2040
Db 1981 TCAAGGAACCTCTGGTGTGAGATGAAAAGAGTTCTGAACTTATTATTACTGATTTCAATAA 2040
QY 2041 CCCTGAAGATAAAAACGGAAATCAAGTCTTCTAGCTAAAATTAAGAACTTAAAGAGTATC 2100
Db 2041 CCCTGAAGATAAAAACGGAAATCAAGTCTTCTAGCTAAAATTAAGAACTTAAAGAGTATC 2100
QY 2101 AAGAACCCAGAGGTTCCAGAGAGTAACCCAGAAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGT 2160
Db 2101 AAGAACCCAGAGGTTCCAGAGAGTAACCCAGAAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGT 2160
QY 2161 GTATTAAACCAGAAATCCTGCTGCTGATCTTTCAAATCACTGGCAGATTTCCGGAGTTAGCCCGAA 2220
Db 2161 GTATTAAACCAGAAATCCTGCTGCTGATCTTTCAAATCACTGGCAGATTTCCGGAGTTAGCCCGAA 2220
QY 2221 AAGTTAATAACAGAGCAGAAAACATACCACCTTTTGGCAACCTGTCTTTTCAGTTTCAAAAAC 2280
Db 2221 AAGTTAATAACAGAGCAGAAAACATACCACCTTTTGGCAACCTGTCTTTTCAGTTTCAAAAAC 2280
QY 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGAACCAAAATCTATTGTAAAGACACCAA 2340
Db 2281 AGTCACCAACCAATATCAACATCTAAATGGTTTGAACCAAAATCTATTGTAAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTGAACCTCCAGTTGTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTAATGAGCTGTTTGAACCTCCAGTTGTAAAGAAATGACT 2400
QY 2401 TTCCACCTGCTTGTGAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGAGTTGTCAACACCTTATGGCCAAACCTGCCTGTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAAAATACTTGCCACTTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAAATACTTGCCACTTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAA 2520
QY 2521 ATGAATGCAATTCGGTTTAAAGGAAGAATTTATTCCATATTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCAATTCGGTTTAAAGGAAGAATTTATTCCATATTAAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGATTTTCAGGTGTTAAATGAAGAAAACAGATATATGCTATAAAAATATG 2640
Db 2581 GTTCAAGCAAGGATTTTCAGGTGTTAAATGAAGAAAACAGATATATGCTATAAAAATATG 2640
QY 2641 TGAACCTTAGAAGAAGCAGATAACCAACTCTTGTAGATTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAAGCAGATAACCAACTCTTGTAGATTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAACAACACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCACGG 2760
Db 2701 TGAATAAACTACAACAACACAGTGATAAGATCATCCGACTTTATGATTTATGAAATCACGG 2760

QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAA 2820
Db |||||
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAA 2820
Db |||||
QY 2821 AGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAATAATATGTTAGAGGCAG 2880
Db |||||
QY 2821 AGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAATAATATGTTAGAGGCAG 2880
Db |||||
QY 2881 TTCACACAATCCATCAACATGGCAATGTTTACAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Db |||||
QY 2881 TTCACACAATCCATCAACATGGCAATGTTTACAGTGATCTTAAACCAGCTAACTTTCTGA 2940
Db |||||
QY 2941 TAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTTGCAAAACCAATGCAACAGATA 3000
Db |||||
QY 2941 TAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTTGCAAAACCAATGCAACAGATA 3000
Db |||||
QY 3001 CAACAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATTATATGCCACCAGAACAA 3060
Db |||||
QY 3001 CAACAAGTGTGTTAAAGATTCTCAGGTGGCACAGTTAATTATATGCCACCAGAACAA 3060
Db |||||
QY 3061 TCAAAGATATGTTCTCTCCAGAGAGAAATCTAAGTCAAAGATAAGCCCCAAAA 3120
Db |||||
QY 3061 TCAAAGATATGTTCTCTCCAGAGAGAAATCTAAGTCAAAGATAAGCCCCAAAA 3120
Db |||||
QY 3121 GTGATGTTTGGTCTTAGGATGTTATTTGTACTATATGACTACGGGAAAAACCATTTTC 3180
Db |||||
QY 3121 GTGATGTTTGGTCTTAGGATGTTATTTGTACTATATGACTACGGGAAAAACCATTTTC 3180
Db |||||
QY 3181 AGCAGATAATTAATCAGATTTCTAAATACATGCCATAATGATCCTAATCATGAAATTG 3240
Db |||||
QY 3181 AGCAGATAATTAATCAGATTTCTAAATACATGCCATAATGATCCTAATCATGAAATTG 3240
Db |||||
QY 3241 AATTTCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db |||||
QY 3241 AATTTCCGATATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db |||||
QY 3301 ACCCAAAACAGAGATATCCATTCTCAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db |||||
QY 3301 ACCCAAAACAGAGATATCCATTCTCAGCTCCTGGCTCATCCATATGTTCAAATTCAAA 3360
Db |||||
QY 3361 CTCATCCAGTTAACCAATGGCCAAAGCAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
Db |||||
QY 3361 CTCATCCAGTTAACCAATGGCCAAAGCAACCACTGAAGAAATGAAATATGTTCTGGGCC 3420
Db |||||
QY 3421 AACTTGTGTTCTGAATTCCTCACTCCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db |||||
QY 3421 AACTTGTGTTCTGAATTCCTCACTCCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
Db |||||
QY 3481 ACTATAGTGGTGAAGTCAATATCTTCTCATCTCCTCAAGCTGCTGCTGCTGCTGCTGCTGCT 3540
Db |||||
QY 3481 ACTATAGTGGTGAAGTCAATATCTTCTCATCTCCTCAAGCTGCTGCTGCTGCTGCTGCTGCT 3540
Db |||||
QY 3541 GAAAAAATGATTTGCGAGTTATTCGTAATGTCAGATAGGAGGTATATAATATATTGGACT 3600
Db |||||
QY 3541 GAAAAAATGATTTGCGAGTTATTCGTAATGTCAGATAGGAGGTATATAATATATTGGACT 3600
Db |||||
QY 3601 GTTATACCTTGAATCCCTGTGGAAATCTACATTTGAAGACACATCACTCTGAAGTGT 3660
Db |||||
QY 3601 GTTATACCTTGAATCCCTGTGGAAATCTACATTTGAAGACACATCACTCTGAAGTGT 3660
Db |||||
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAACTGTAAAAATAGCAACCAC 3720
Db |||||
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAACTGTAAAAATAGCAACCAC 3720
Db |||||
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTGTTTCTCTGTTTATGCTCTTGTGTAACTAC 3780
Db |||||
QY 3721 TTATGGCACTGTATATATTGTAGACTTGTGTTTCTCTGTTTATGCTCTTGTGTAACTAC 3780
Db |||||
QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGGCGATAGCAAGTATATTCTAAAAAACTTG 3840
Db |||||
QY 3781 TTGACATCATTTTACTCTTGGAAATAGTGGGCGATAGCAAGTATATTCTAAAAAACTTG 3840
Db |||||
QY 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866

Db 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866
|||
RESULT 6
ADE38368
ID ADE38368 standard; DNA; 3866 BP.
XX
AC ADE38368;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human protein 1682 gene sequence.
XX
KW tumorigenic disorder; angiogenic disorder; aberrant gene expression;
KW aberrant protein activity; cytostatic; antithyroid; antidiabetic;
KW ophthalmological; cancer; breast cancer; colon cancer; lung cancer;
KW prostatic cancer; Grave's disease; diabetic retinopathy; gene; ds;
KW protein 1682.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 978..3551
FT /*tag= a
FT /product= "Human protein 1682"
XX
PN WC2003065006-A2.
XX
PD 07-AUG-2003.
XX
PF 30-JAN-2003; 2003WO-US002588.
XX
PR 31-JAN-2002; 2002US-0353600P.
PR 15-MAR-2002; 2002US-0364517P.
PR 09-APR-2002; 2002US-0371075P.
PR 10-APR-2002; 2002US-0371507P.
PR 16-APR-2002; 2002US-0372984P.
PR 19-APR-2002; 2002US-0374194P.
PR 24-MAY-2002; 2002US-0382995P.
PR 31-MAY-2002; 2002US-0385023P.
PR 14-JUN-2002; 2002US-0388533P.
PR 17-JUN-2002; 2002US-0389395P.
PR 25-JUN-2002; 2002US-0391324P.
PR 15-JUL-2002; 2002US-0395944P.
PR 22-JUL-2002; 2002US-0397726P.
PR 13-AUG-2002; 2002US-0403046P.
PR 22-AUG-2002; 2002US-0405155P.
PR 27-AUG-2002; 2002US-0406361P.
PR 25-OCT-2002; 2002US-0421195P.
PR 12-NOV-2002; 2002US-0425456P.
PR 19-NOV-2002; 2002US-0427626P.
PR 10-DEC-2002; 2002US-0432122P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Hunter JJ, Macbeth KJ, Tsai F, Lesoon A, Lightcap ES;
PI Williamson MW, Rudolph-Owen LA;
XX
DR WPI; 2003-646176/61.
DR P-PSDB; ADE38369.
XX
PT Treating subject having tumorigenic disorder or angiogenic disorder
PT caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic
PT acid, by administering a modulator.
XX
PS Disclosure; SEQ ID NO 29; 454pp; English.
XX
CC This invention relates to a novel method of treating a human subject
CC having a tumorigenic disorder or angiogenic disorder, caused by aberrant
CC gene expression or activity of an isolated protein, by administering a
CC modulator. The modulator may have cytostatic, antithyroid, antidiabetic
CC or ophthalmological activity. The method is useful for treating a subject

CC having a tumorigenic or angiogenic disorder, in particular for treating
CC cancer (for example breast cancer, colon cancer, lung cancer or prostatic
CC cancer) and, for example, Grave's disease and diabetic retinopathy. The
CC present sequence is a DNA sequence which encodes the novel isolated human
CC protein 1682 of the invention.
XX

SQ Sequence 3866 BP; 1268 A; 731 C; 752 G; 1115 T; 0 U; 0 Other;

Query Match 100.0%; Score 3866; DB 9; Length 3866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3866; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGAAATCCCTTTT	1	GGAAATCCCTTTT	60
DB	1	GGAAATCCCTTTT	1	GGAAATCCCTTTT	60
QY	61	CAATGGCAAACTC	61	CAATGGCAAACTC	120
DB	61	CAATGGCAAACTC	61	CAATGGCAAACTC	120
QY	121	CTCAGCCTCTCA	121	CTCAGCCTCTCA	180
DB	121	CTCAGCCTCTCA	121	CTCAGCCTCTCA	180
QY	181	CTTTTCTATTAG	181	CTTTTCTATTAG	240
DB	181	CTTTTCTATTAG	181	CTTTTCTATTAG	240
QY	241	ACCTCAGGTGAT	241	ACCTCAGGTGAT	300
DB	241	ACCTCAGGTGAT	241	ACCTCAGGTGAT	300
QY	301	TGCCTGGCTGAT	301	TGCCTGGCTGAT	360
DB	301	TGCCTGGCTGAT	301	TGCCTGGCTGAT	360
QY	361	AGGCTGGAGTGC	361	AGGCTGGAGTGC	420
DB	361	AGGCTGGAGTGC	361	AGGCTGGAGTGC	420
QY	421	GATCCTCCCACT	421	GATCCTCCCACT	480
DB	421	GATCCTCCCACT	421	GATCCTCCCACT	480
QY	481	CTAATTTTGTAT	481	CTAATTTTGTAT	540
DB	481	CTAATTTTGTAT	481	CTAATTTTGTAT	540
QY	541	CTCCTGGACTCA	541	CTCCTGGACTCA	600
DB	541	CTCCTGGACTCA	541	CTCCTGGACTCA	600
QY	601	AGTCACCATGCC	601	AGTCACCATGCC	660
DB	601	AGTCACCATGCC	601	AGTCACCATGCC	660
QY	661	GGTCCTCACAAC	661	GGTCCTCACAAC	720
DB	661	GGTCCTCACAAC	661	GGTCCTCACAAC	720
QY	721	AGGAAGATGAAG	721	AGGAAGATGAAG	780
DB	721	AGGAAGATGAAG	721	AGGAAGATGAAG	780
QY	781	GGATGATGATCC	781	GGATGATGATCC	840
DB	781	GGATGATGATCC	781	GGATGATGATCC	840
QY	841	TTTGGTTTAAAT	841	TTTGGTTTAAAT	900
DB	841	TTTGGTTTAAAT	841	TTTGGTTTAAAT	900
QY	901	ATGTAAGACTGT	901	ATGTAAGACTGT	960

DB	901	ATGTAAGACTGT	901	ATGTAAGACTGT	960
QY	961	TGCAGTTTCTGT	961	TGCAGTTTCTGT	1020
DB	961	TGCAGTTTCTGT	961	TGCAGTTTCTGT	1020
QY	1021	CCATAATGAACA	1021	CCATAATGAACA	1080
DB	1021	CCATAATGAACA	1021	CCATAATGAACA	1080
QY	1081	AACTAAGCTTGA	1081	AACTAAGCTTGA	1140
DB	1081	AACTAAGCTTGA	1081	AACTAAGCTTGA	1140
QY	1141	TTATGATGATGG	1141	TTATGATGATGG	1200
DB	1141	TTATGATGATGG	1141	TTATGATGATGG	1200
QY	1201	ACAGTGTCCGCT	1201	ACAGTGTCCGCT	1260
DB	1201	ACAGTGTCCGCT	1201	ACAGTGTCCGCT	1260
QY	1261	TTGAAGCGCTTC	1261	TTGAAGCGCTTC	1320
DB	1261	TTGAAGCGCTTC	1261	TTGAAGCGCTTC	1320
QY	1321	GATTTGCTGAAT	1321	GATTTGCTGAAT	1380
DB	1321	GATTTGCTGAAT	1321	GATTTGCTGAAT	1380
QY	1381	CCAGAGCAAACT	1381	CCAGAGCAAACT	1440
DB	1381	CCAGAGCAAACT	1381	CCAGAGCAAACT	1440
QY	1441	CACAAAGTAATG	1441	CACAAAGTAATG	1500
DB	1441	CACAAAGTAATG	1441	CACAAAGTAATG	1500
QY	1501	TACCACTAGAAAT	1501	TACCACTAGAAAT	1560
DB	1501	TACCACTAGAAAT	1501	TACCACTAGAAAT	1560
QY	1561	TTTCAGAGGAGG	1561	TTTCAGAGGAGG	1620
DB	1561	TTTCAGAGGAGG	1561	TTTCAGAGGAGG	1620
QY	1621	TTTCGGGTCAC	1621	TTTCGGGTCAC	1680
DB	1621	TTTCGGGTCAC	1621	TTTCGGGTCAC	1680
QY	1681	CTACTAAAGCC	1681	CTACTAAAGCC	1740
DB	1681	CTACTAAAGCC	1681	CTACTAAAGCC	1740
QY	1741	ACCGGAATTCAT	1741	ACCGGAATTCAT	1800
DB	1741	ACCGGAATTCAT	1741	ACCGGAATTCAT	1800
QY	1801	CAGTAAACCTTC	1801	CAGTAAACCTTC	1860
DB	1801	CAGTAAACCTTC	1801	CAGTAAACCTTC	1860
QY	1861	GTTTTATGAAA	1861	GTTTTATGAAA	1920
DB	1861	GTTTTATGAAA	1861	GTTTTATGAAA	1920
QY	1921	AACCAAGTGGAA	1921	AACCAAGTGGAA	1980
DB	1921	AACCAAGTGGAA	1921	AACCAAGTGGAA	1980
QY	1981	TCAAGGAACCTC	1981	TCAAGGAACCTC	2040

Db 1981 TCAAGGAACCTTGGTGTCCAGATGAAAAAGAGTTCTGAACTTATTATTACTGATTCATAA 2040
QY 2041 CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAAAACCTAAGAGTATC 2100
Db 2041 CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAAAACCTAAGAGTATC 2100
QY 2101 AAGAACCCAGAGTTCCAGAGAGTAAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
Db 2101 AAGAACCCAGAGTTCCAGAGAGTAAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCAGAGT 2160
QY 2161 GTATTAAACCAAGATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGATTAGCCCGAA 2220
Db 2161 GTATTAAACCAAGATCCTGCTGCATCTTCAAAATCACTGGCAGATTCGGAGATTAGCCCGAA 2220
QY 2221 AAGTTAATACAGAGCAGAAAACATACCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
Db 2221 AAGTTAATACAGAGCAGAAAACATACCACTTTTGAGCAACCTGTCTTTTCAGTTTCAAAAC 2280
QY 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA 2340
Db 2281 AGTCACCACCAATATCAACATCTAAATGGTTTGACCCAAAATCTATTGTGAAGACACCAA 2340
QY 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
Db 2341 GCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACT 2400
QY 2401 TTCCACCTGCTTGTGTCAGTTGTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
Db 2401 TTCCACCTGCTTGTGTCAGTTGTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAAC 2460
QY 2461 AGCATCAAATACCTGCCACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCAGCAA 2520
Db 2461 AGCATCAAATACCTGCCACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCAGCAA 2520
QY 2521 ATGAATGCATTTTCGGTTAAAGGAAGATTTATCCATATTTAAAGCAGATAGGAAGTGGAG 2580
Db 2521 ATGAATGCATTTTCGGTTAAAGGAAGATTTATCCATATTTAAAGCAGATAGGAAGTGGAG 2580
QY 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
Db 2581 GTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAATATG 2640
QY 2641 TGAACCTTAGAAGAACGAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
Db 2641 TGAACCTTAGAAGAACGAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATT 2700
QY 2701 TGAATAAACTACAAACAACAGTGATAAGATCATCCGACTTTATGATTAAGAAATCACGG 2760
Db 2701 TGAATAAACTACAAACAACAGTGATAAGATCATCCGACTTTATGATTAAGAAATCACGG 2760
QY 2761 ACCAGTACATCTACATGGTAATGGAGTGGGAAATATTGATCTTAATAGTTGGCTTAAAA 2820
Db 2761 ACCAGTACATCTACATGGTAATGGAGTGGGAAATATTGATCTTAATAGTTGGCTTAAAA 2820
QY 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTTACTGGAATAATATGTTAGAGGCAG 2880
Db 2821 AGAAAAAATCCATTGATCCATGGGAACGCAAGAGTTTACTGGAATAATATGTTAGAGGCAG 2880
QY 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAAGCTAATCTTCTGA 2940
Db 2881 TTCACACAATCCATCAACATGGCATTTGTTACAGTGATCTTAAACCAAGCTAATCTTCTGA 2940
QY 2941 TAGTTGATGAATGCTAAAAGCTAATTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGATA 3000
Db 2941 TAGTTGATGAATGCTAAAAGCTAATTGATTTTGGGATTGCAAAACCAAAATGCAACCCAGATA 3000
QY 3001 CAACAAGTGTGTTTAAAGATTCTCAGGTTGGCAGATTAAATTATATGCCACCAGAGACAA 3060
Db 3001 CAACAAGTGTGTTTAAAGATTCTCAGGTTGGCAGATTAAATTATATGCCACCAGAGACAA 3060
QY 3061 TCAAGAATATGTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCCAAAA 3120
Db 3061 TCAAGAATATGTCTTCTCCAGAGAGAAATGGGAAATCTAAGTCAAAGATAAGCCCCAAAA 3120

QY 3121 GTGATGTTTGGTCTTAGGATGTATTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
Db 3121 GTGATGTTTGGTCTTAGGATGTATTTTGTACTATATGACTTACGGGAAAAACACCATTTTC 3180
QY 3181 AGCAGATAAATTAAATCAGATTTCCTAAATTAACATGCCATAAATTGATCCTTAATCATGAAATG 3240
Db 3181 AGCAGATAAATTAAATCAGATTTCCTAAATTAACATGCCATAAATTGATCCTTAATCATGAAATG 3240
QY 3241 AATTTCCCGATATTCAGAGAAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
Db 3241 AATTTCCCGATATTCAGAGAAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGG 3300
QY 3301 ACCCAAAAACAGAGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTTCAAAATTCAAA 3360
Db 3301 ACCCAAAAACAGAGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTTCAAAATTCAAA 3360
QY 3361 CTCATCCAGTTAACCAAAATGGCCAAGGAACCACTGAAGAAATGAATAATATGTTCTGGGCC 3420
Db 3361 CTCATCCAGTTAACCAAAATGGCCAAGGAACCACTGAAGAAATGAATAATATGTTCTGGGCC 3420
QY 3421 AACTTGTGGTCTGAATTCCTTAACCTCACTTTTGAAGCTGCTAAAACTTTTATATGAAC 3480
Db 3421 AACTTGTGGTCTGAATTCCTTAACCTCACTTTTGAAGCTGCTAAAACTTTTATATGAAC 3480
QY 3481 ACTATAGTGGTGGTGAAGTCATAAATCTCTCATCTCCTCCAAGACTTTTGAAGAAAAAGGG 3540
Db 3481 ACTATAGTGGTGGTGAAGTCATAAATCTCTCATCTCCTCCAAGACTTTTGAAGAAAAAGGG 3540
QY 3541 GAAAAAAATGATTTCGAGTTATTCGTAATGTCAGATAGGAGGTATAAAATATATTTGGACT 3600
Db 3541 GAAAAAAATGATTTCGAGTTATTCGTAATGTCAGATAGGAGGTATAAAATATATTTGGACT 3600
QY 3601 GTTATACCTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
Db 3601 GTTATACCTCTTGAATCCCTGTGGAAATCTACATTTGAAGACAACATCACTCTGAAGTGT 3660
QY 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAACTGTAAAAAATAGCAACCAC 3720
Db 3661 ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAAAGAAAACTGTAAAAAATAGCAACCAC 3720
QY 3721 TTATGGCACTGTATATATTTGAGACTTGTGTTTCTCTGTTTTATGCTCTTGTGTAATCTAC 3780
Db 3721 TTATGGCACTGTATATATTTGAGACTTGTGTTTCTCTGTTTTATGCTCTTGTGTAATCTAC 3780
QY 3781 TTGACATCATTTTACTCTTGAATAGTGGTGGATAGCAAGTATATTTAAAAAACTTTT 3840
Db 3781 TTGACATCATTTTACTCTTGAATAGTGGTGGATAGCAAGTATATTTAAAAAACTTTT 3840
QY 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866
Db 3841 TAAATAAAGTTTGTGGCTAAAAATGA 3866

RESULT 7
ADE31644
ID ADE31644 standard; DNA; 3866 BP.
XX ADE31644;
XX AC
XX DT 29-JAN-2004 (first entry)
XX DE Human 1682 gene #SEQ ID 1.
XX KW Antiartherosclerotic; cardiant; vasotropic; antiinflammatory;
KW thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;
KW cardiovascular; disorder; ischaemia; aortic bending;
KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;
KW angina; cardiomyopathy; cardiac death; gene; ds.
XX OS Homo sapiens.
XX PN W02003065984-A2.

Db	1141	TTATGATGATGGCAAAACAACCCAGAGGAGTGGTTGAGTTTGTTCCTCAAACTAGAGAAA	1200
Qy	1201	ACAGTGTTCGGCTAAGTGATGCTCTTTTAAATAAATTGATTGGTCGTTACAGTCAAGCAA	1260
Db	1201	ACAGTGTTCGGCTAAGTGATGCTCTTTTAAATAAATTGATTGGTCGTTACAGTCAAGCAA	1260
Qy	1261	TTGAAGCGTTCCTCCCCAGATAAAATATGGCCAAAAATGAGAGTTTTTGTCTAGAATTCAAAGTGA	1320
Db	1261	TTGAAGCGTTCCTCCCCAGATAAAATATGGCCAAAAATGAGAGTTTTTGTCTAGAATTCAAAGTGA	1320
Qy	1321	GATTTGCTGAATTTAAAAGCTATTCAAGAGCCAGATGATGCAAGTACTATCTTTTCAAATGG	1380
Db	1321	GATTTGCTGAATTTAAAAGCTATTCAAGAGCCAGATGATGCAAGTACTATCTTTTCAAATGG	1380
Qy	1381	CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTGAACTGT	1440
Db	1381	CCAGAGCAAACTGCAAGAAAATTTGCTTTTGTTCATATATCTTTTGCACAATTTGAACTGT	1440
Qy	1441	CACAAAGTAAATGTCAAAAAAAGTAAACAACCTCTTCAAAAAAGCTGTAGAACGTGGAGCAG	1500
Db	1441	CACAAAGTAAATGTCAAAAAAAGTAAACAACCTCTTCAAAAAAGCTGTAGAACGTGGAGCAG	1500
Qy	1501	TACCACTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC	1560
Db	1501	TACCACTAGAAATGCTGGAAATTTGCCCTGCGGAATTTAAACCTCCAAAAAAGCAGCTGC	1560
Qy	1561	TTTCAGAGGAGAAAAAGAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT	1620
Db	1561	TTTCAGAGGAGAAAAAGAAGAAATTTATCAGCATCTACGGTATTAACTGCCCAAGAAATCAT	1620
Qy	1621	TTTCCGGTTCACTTGGGCATTTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA	1680
Db	1621	TTTCCGGTTCACTTGGGCATTTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGA	1680
Qy	1681	CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCACAAGATGCAGAAATAGGTT	1740
Db	1681	CTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCACAAGATGCAGAAATAGGTT	1740
Qy	1741	ACCGGAATTCATTGAGACAAACTRAACAAACTAAACAGTCATGCCCATTTTGGAAAGAGTCC	1800
Db	1741	ACCGGAATTCATTGAGACAAACTRAACAAACTAAACAGTCATGCCCATTTTGGAAAGAGTCC	1800
Qy	1801	CAGTTAACCTTCTAAATAGCCCCAGATTGTGATGTGAAGACAGATGATTTCAGTTGTACCTT	1860
Db	1801	CAGTTAACCTTCTAAATAGCCCCAGATTGTGATGTGAAGACAGATGATTTCAGTTGTACCTT	1860
Qy	1861	GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCTGGATCTA	1920
Db	1861	GTTTTATGAAAAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCCTGGATCTA	1920
Qy	1921	AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAATAAGTAGTCATT	1980
Db	1921	AACCAAGTGGAAATGATTCCTGTGAAATTAAGAAATTTAAAGTCTGTTCAAATAAGTAGTCATT	1980
Qy	1981	TCAAGGAACCTCTGGTGTCCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCATATA	2040
Db	1981	TCAAGGAACCTCTGGTGTCCAGATGAAAAGAGTTCTGAACCTTATTATTACTGATTCATATA	2040
Qy	2041	CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC	2100
Db	2041	CCCTGAAGAATAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATC	2100
Qy	2101	AAGAACCCAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGT	2160
Db	2101	AAGAACCCAGAGGTTCCAGAGAGTAACAGAAACAGTGGCAAGCTAAGAGAAAAGTCAGAGT	2160
Qy	2161	GTATTAAACCAGAAATCCTGTGTCATCTTCAAATCACTGGCAGATTCGGAGTTAGCCCGAA	2220
Db	2161	GTATTAAACCAGAAATCCTGTGTCATCTTCAAATCACTGGCAGATTCGGAGTTAGCCCGAA	2220
Qy	2221	AAGTTAATACAGAGCAGAAAAACATACCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAC	2280
Db	2221	AAGTTAATACAGAGCAGAAAAACATACCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAAC	2280

QY 1705 GAGAGAACATGCCACCACAAGATGCAGAAATAGGTTACCGGAATTCATTGGAGACAAACTA 1764
Dbb 811 GAGAGAACATGCCACCACAAGATGCAGAAATAGGTTACCGGAATTCATTGGAGACAAACTA 870
QY 1765 ACAAAACTAAACAGTCATGCCCATTTGGAAAGAGTCCCGATTAACTTCTTAATAGCCCGAG 1824
Dbb 871 ACAAAACTAAACAGTCATGCCCATTTGGAAAGAGTCCCGATTAACTTCTTAATAGCCCGAG 930
QY 1825 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTACCTTTATGAAAAGACAACTCTTA 1884
Dbb 931 ATTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTATGAAAAGACAACTCTCTA 990
QY 1885 GATCAGAAATGCCAGATTTGGTTGTGCTGGATCTAAACCAGTGGAAATGATTCCTGTG 1944
Dbb 991 GATCAGAAATGCCAGATTTGGTTGTGCTGGATCTAAACCAGTGGAAATGATTCCTGTG 1050
QY 1945 AATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCATTTCAAGAAACCTCTGTGTGATG 2004
Dbb 1051 AATTAAGAAATTTAAAGTCTGTTTCAAAATAGTCATTTCAAGAAACCTCTGTGTGATG 1110
QY 2005 AAAAGAGTTCTGAACCTTATTATTACTGATTCATTAACCTCGAAGAAATAAAACCGAATCAA 2064
Dbb 1111 AAAAGAGTTCTGAACCTTATTATTACTGATTCATTAACCTCGAAGAAATAAAACCGAATCAA 1170
QY 2065 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACACAGAGTTCCAGAGAGTA 2124
Dbb 1171 GTCTTCTAGCTAAATTAAGAAACTAAAGAGTATCAAGAACACAGAGTTCCAGAGAGTA 1230
QY 2125 ACCAGAAACAGTGCAGCTAAGAGAAAGTCAAGAGTGTATTAAACAGAAATCCTGCTGCAT 2184
Dbb 1231 ACCAGAAACAGTGCAGCTAAGAGAAAGTCAAGAGTGTATTAAACAGAAATCCTGCTGCAT 1290
QY 2185 CTTCAAATCACTGCGAGATTCGGGAGTTAGCCCGGAAAGTTAATACAGAGCAGAGAAACATA 2244
Dbb 1291 CTTCAAATCACTGCGAGATTCGGGAGTTAGCCCGGAAAGTTAATACAGAGCAGAGAAACATA 1350
QY 2245 CCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCACCACCAATATCAACATCTA 2304
Dbb 1351 CCACCTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCACCACCAATATCAACATCTA 1410
QY 2305 AATGGTTTGACCCCAAAATCTATTTGTAAACACCAAGCAGCAATACCTTGGATGATTACA 2364
Dbb 1411 AATGGTTTGACCCCAAAATCTATTTGTAAACACCAAGCAGCAATACCTTGGATGATTACA 1470
QY 2365 TGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTGATGTCATA 2424
Dbb 1471 TGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGCTTGTGATGTCATA 1530
QY 2425 CACCTTATGGCCCAACCTGCTGTTTCCAGCAGCAACAGCATCAAAATACTTGCACCTCCAC 2484
Dbb 1531 CACCTTATGGCCCAACCTGCTGTTTCCAGCAGCAACAGCATCAAAATACTTGCACCTCCAC 1590
QY 2485 TTCAAAATTTACAGGTTTTAGCATCTTCTTCAGCAAAATGAATGCAATTCGGTTTAAAGGAA 2544
Dbb 1591 TTCAAAATTTACAGGTTTTAGCATCTTCTTCAGCAAAATGAATGCAATTCGGTTTAAAGGAA 1650
QY 2545 GAATTTATTCCATATTAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTCAGGTGT 2604
Dbb 1651 GAATTTATTCCATTTAAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGTATTTTCAGGTGT 1710
QY 2605 TAAATGAAAAGAAACAGATATATGCTATATAAATATGTGAACCTTAGAAGAACAGATAAACC 2664
Dbb 1711 TAAATGAAAAGAAACAGATATATGCTATATAAATATGTGAACCTTAGAAGAACAGATAAACC 1770
QY 2665 AAACTCTTGATAGTTACCGGAACGAAATAGCTTATTGTAATAAATCAACACACAGTG 2724
Dbb 1771 AAACTCTTGATAGTTACCGGAACGAAATAGCTTATTGTAATAAATCAACACACAGTG 1830
QY 2725 ATAAGATCATCCGACITTTATGATTATGAATCAGGACCGACAGTACATCTACATGGTAATGG 2784
Dbb 1831 ATAAGATCATCCGACITTTATGATTATGAATCAGGACCGACAGTACATCTACATGGTAATGG 1890
QY 2785 AGTGTGGAATATTGATCTTAAATAGTTGGCTTAAAGAAAGAAATCCATTGATCCATGGG 2844

Dbb 1891 AGTGTGGAATATTGATCTTAAATAGTTGGCTTAAAAAGAAAAATCCATTGATCCATGGG 1950
QY 2845 AACGCAAGAGTTACTGGAATAATATGTAGAGGCAGTTACACAAATCCATCAACATGGCA 2904
Dbb 1951 AACGCAAGAGTTACTGGAATAATATGTAGAGGCAGTTACACAAATCCATCAACATGGCA 2010
QY 2905 TTGTTACAGTGATCTTAAACCAGCTAACTTTCTGATAGTTGATGGAATGCTAAAGCTAA 2964
Dbb 2011 TTGTTACAGTGATCTTAAACCAGCTAACTTTCTGATAGTTGATGGAATGCTAAAGCTAA 2070
QY 2965 TTGATTTTGGGATTGCAAAACCAAAATGCAACCAGATACAAACAGTGTGTTTAAAGATTCTC 3024
Dbb 2071 TTGATTTTGGGATTGCAAAACCAAAATGCAACCAGATACAAACAGTGTGTTTAAAGATTCTC 2130
QY 3025 AGGTTGGCACAGTTAATTATATGCCACCAGAAACCAATCAAAAGATATGTTCTTCCAGAG 3084
Dbb 2131 AGGTTGGCACAGTTAATTATATGCCACCAGAAACCAATCAAAAGATATGTTCTTCCAGAG 2190
QY 3085 AGAATGGGAATCTTAAGTCAAGATAAGCCCCAAAAGTGTGTTGGTCTCTTAGGATGTA 3144
Dbb 2191 AGAATGGGAATCTTAAGTCAAGATAAGCCCCAAAAGTGTGTTGGTCTCTTAGGATGTA 2250
QY 3145 TTTTGTACTATATGACTTACGGGAAAACACCATTTTCAGCAGATATTAATCAGATTCTTA 3204
Dbb 2251 TTTTGTACTATATGACTTACGGGAAAACACCATTTTCAGCAGATATTAATCAGATTCTTA 2310
QY 3205 AATTACATGCCATTAATTGATCCTTAATCATGAATAATGAAATTTCCCGATATTCAGAGAAAG 3264
Dbb 2311 AATTACATGCCATTAATTGATCCTTAATCATGAATAATGAAATTTCCCGATATTCAGAGAAAG 2370
QY 3265 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGGACCCAAAACAGAGGATATCCATTTC 3324
Dbb 2371 ATCTTCAAGATGTGTTAAAGTGTGTTTAAAAAGGGACCCAAAACAGAGGATATCCATTTC 2430
QY 3325 CTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAACTCATCCAGTTAAACCAATGGCCA 3384
Dbb 2431 CTGAGCTCCTGGCTCATCCCTATGTTCAAATTCAAACTCATCCAGTTAAACCAATGGCCA 2490
QY 3385 AGGAAACCACTGAAGAAATGAAATATGTTCTGGGCCAACTTGTGTTGTTCTGAAATTCCTTA 3444
Dbb 2491 AGGAAACCACTGAAGAAATGAAATATGTTCTGGGCCAACTTGTGTTGTTCTGAAATTCCTTA 2550
QY 3445 ACTCCATTTTGAAGCTGCTAAAACTTTATATGAACACTATAGTGGTGGTGAAGTCTATA 3504
Dbb 2551 ACTCCATTTTGAAGCTGCTAAAACTTTATATGAACACTATAGTGGTGGTGAAGTCTATA 2610
QY 3505 ATTCTTTCATCCTCCAGACTTTTGAATAAAAGGGGAAATAATGATTTGCAGTTATTC 3564
Dbb 2611 ATTCTTTCATCCTCCAGACTTTTGAATAAAAGGGGAAATAATGATTTGCAGTTATTC 2670
QY 3565 GTAAATGTCAGATAGGAGGTATAAAATATATTGGACTGTTTATACCTCTTGAATCCCTGTGGA 3624
Dbb 2671 GTAAATGTCAAATACCACTTATAAAATATATTGGACTGTTTATACCTCTTGAATCCCTGTGGA 2730
QY 3625 AATCTACATTTGAAGACAAACATCACTCTGAAAGTGTATTATCAGCAAAAAAATTCAGTGGA 3684
Dbb 2731 AATCTACATTTGAAGACAAACATCACTCTGAAAGTGTATTATCAGCAAAAAAATTCAGT - AGA 2789
QY 3685 TTATCTTTAAAGAAACTGTAAAAATAGCAACCACTTATGGCAGTGTATATATTGTAGA 3744
Dbb 2790 TTATCTTTAAAGAAACTGTAAAAATAGCAACCACTTATGGTACTGTATATATTGTAGA 2849
QY 3745 CTTGTTTTCTCTGTTTTATGCTCTTGTGTAATCTACTTGCATCATTTTACTCTTTGGAAT 3804
Dbb 2850 CTTGTTTTCTCTGTTTTATGCTCTTGTGTAATCTACTTGCATCATTTTACTCTTTGGAAT 2909
QY 3805 AGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTGTAAATAAAGTTTGTGGCTAAAAAT 3864
Dbb 2910 AGTGGGTGGATAGCAAGTATATTCTAAAAAACTTTGTAAATAAAGTTTGTGGCTAAAAAT 2969
QY 3865 GA 3866
||

Db 781 CAGACTACTAAAGCCAGGTTTTTATATGGAGAGAAACATGCCACCAACAGATGCAGAAATA 840

Qy 1737 GGTACCGGAATTCATTGAGACAAAACHTAACAAAACTAAACAGTCATGCCCATTTGGAAGA 1796

Db 841 GGTACCGGAATTCATTGAGACAAAACHTAACAAAACTAAACAGTCATGCCCATTTGGAAGA 900

Qy 1797 GTCCCCAGTTAACCTTCTAAATAGCCCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTA 1856

Db 901 GTCCCCAGTTAACCTTCTAAATAGCCCCAGATTGTGATGTGAAGACAGATGATTCAGTTGTA 960

Qy 1857 CCTTGTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGA 1916

Db 961 CCTTGTTTTATGAAAAGACAAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGA 1020

Qy 1917 TCTAAACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAATAGT 1976

Db 1021 TCTAAACCAAGTGGAAATGATTCCTGTGAATTAAGAAATTTAAAGTCTGTTCAAATAGT 1080

Qy 1977 CATTTCAAGGAACCTCTGGTGTCCAGATGAAAAAGAGTTCTGAACTTATTATTACTGATTCA 2036

Db 1081 CATTTCAAGGAACCTCTGGTGTCCAGATGAAAAAGAGTTCTGAACTTATTATTACTGATTCA 1140

Qy 2037 ATAAACCTGAAGATAAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAG 2096

Db 1141 ATAAACCTGAAGATAAAAACGGAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAG 1200

Qy 2097 TATCAAGAACCCAGAGGTTCCAGAGAGTAAACAGAAAACAGTGGCAAGCTAAGAGAAAGTCA 2156

Db 1201 TATCAAGAACCCAGAGGTTCCAGAGAGTAAACAGAAAACAGTGGCAATCTAAGAGAAAGTCA 1260

Qy 2157 GAGTGTATTAACCCAGAAATCCTGCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCC 2216

Db 1261 GAGTGTATTAACCCAGAAATCCTGCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCC 1320

Qy 2217 CGAAAAGTTAATACAGAGCAGAAAACATACCACTTTTGAGCAACCTGTCTTTTCACTTTCA 2276

Db 1321 CGAAAAGTTAATACAGAGCAGAAAACATACCACTTTTGAGCAACCTGTCTTTTCACTTTCA 1380

Qy 2277 AAACAGTCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTTGTAAGACA 2336

Db 1381 AAACAGTCAACCAATATCAACATCTAAATGGTTTGACCCAAATCTATTTGTAAGACA 1440

Qy 2337 CCAAGCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAAT 2396

Db 1441 CCAAGCAGCAATACCTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAGAAT 1500

Qy 2397 GACTTTCACCTGCTGTGAGTTGTCAACACCTTATGGCCACCTGCCTGTTTCCAGCAG 2456

Db 1501 GACTTTCACCTGCTGTGAGTTGTCAACACCTTATGGCCACCTGCCTGTTTCCAGCAG 1560

Qy 2457 CAACAGCATCAAAATFACTTGGCACTCCACTCAAATTTTACAGGTTTGTAGCATCTTCTTCA 2516

Db 1561 CAACAGCATCAAAATFACTTGGCACTCCACTCAAATTTTACAGGTTTGTAGCATCTTCTTCA 1620

Qy 2517 GCAAAATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGT 2576

Db 1621 GCAAAATGAATGCATTTCCGTTAAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGT 1680

Qy 2577 GGAGGTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAA 2636

Db 1681 GGAGGTTCAAGCAAGGTATTTTCAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAA 1740

Qy 2637 TATGTGAACCTTAGAAGAACAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCT 2696

Db 1741 TATGTGAACCTTAGAAGAACAGATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCT 1800

Qy 2697 TATTTGAATAAACTACAACAACAGATGATAAGATCATCCGACTTTATGATTAAGAAATC 2756

Db 1801 TATTTGAATAAACTACAACAACAGATGATAAGATCATCCGACTTTATGATTAAGAAATC 1860

Qy 2757 ACGGACCAAGTACATCTACATGGTAATGGAGTGTGGAATATGATCTTAATAGTTGGCTT 2816

Db 1861 ACGGACCAAGTACATCTACATGGTAATGGAGTGTGGAATATGATCTTAATAGTTGGCTT 1920

Qy 2817 AAAAAAGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAG 2876

Db 1921 AAAAAAGAAAAATCCATTGATCCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAG 1980

Qy 2877 GCAGTTCACACAATCCATCAACATGGCATTGTTACAGTGATCTTAAACCAGCTAACTTT 2936

Db 1981 GCAGTTCACACAATCCATCAACATGGCATTGTTACAGTGATCTTAAACCAGCTAACTTT 2040

Qy 2937 CTGATAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTGCAAAACCAATGCAACCA 2996

Db 2041 CTGATAGTTGATGGAATGCTAAAGCTAAATGATTTTGGGATTGCAAAACCAATGCAACCA 2100

Qy 2997 GATACAAACAAGTGTGTTAAAGATTCTCAGTTGGCACAGTTAATTATATGCCACCAGAA 3056

Db 2101 GATACAAACAAGTGTGTTAAAGATTCTCAGTTGGCACAGTTAATTATATGCCACCAGAA 2160

Qy 3057 GCAATCAAAAGATATGTCCTTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCC 3116

Db 2161 GCAATCAAAAGATATGTCCTTCCAGAGAGAAATGGGAAATCTAAGTCAAAAGATAAGCCCC 2220

Qy 3117 AAAAGTGATGTTTGGTCCTTAGGATGTAATTTGTACTATATGACTTACGGGAAAAACACCA 3176

Db 2221 AAAAGTGATGTTTGGTCCTTAGGATGTAATTTGTACTATATGACTTACGGGAAAAACACCA 2280

Qy 3177 TTTCAGCAGATAATTAATCAGATTCTTAATTTACATGCCATAATTTGATCCTAATCATGAA 3236

Db 2281 TTTCAGCAGATAATTAATCAGATTCTTAATTTACATGCCATAATTTGATCCTAATCATGAA 2340

Qy 3237 ATTGAATTTCCCGATATTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAA 3296

Db 2341 ATTGAATTTCCCGATATTCCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAA 2400

Qy 3297 AGGACCCCAAAACAGAGGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAAT 3356

Db 2401 AGGACCCCAAAACAGAGGATATCCATTCTCTGAGCTCCTGGCTCATCCATATGTTCAAAT 2460

Qy 3357 CAAACTCATCCAGTTTAAACCAATGGCCAAGGAAACCACTGAAGAAATGAAATATGTTCTG 3416

Db 2461 CAAACTCATCCAGTTTAAACCAATGGCCAAGGAAACCACTGAAGAAATGAAATATGTTCTG 2520

Qy 3417 GGCCAACTTGTGTTGCTGAATTTCTCTAACTCCATTTTGAAGCTGTCTAAACTTTTATAT 3476

Db 2521 GGCCAACTTGTGTTGCTGAATTTCTCTAACTCCATTTTGAAGCTGTCTAAACTTTTATAT 2580

Qy 3477 GAACACTATAGTGGTGAAGTCAATAATCTTTCATCCTCCAAGACTTTTGAAAAAAA 3536

Db 2581 GAACACTATAGTGGTGAAGTCAATAATCTTTCATCCTCCAAGACTTTTGAAAAAAA 2640

Qy 3537 AGGGGAAAAAATGATTTGCAGTTATTCTGTAATGTCCAGATAGGAGTATAAAATATATTG 3596

Db 2641 AGGGGAAAAAATGATTTGCAGTTATTCTGTAATGTCCAGATAGGAGTATAAAATATATTG 2700

Qy 3597 GACTGTTATACCTTTGAATCCCTGTGGAATCTACATTTGAAGACAAACATCCTCTGAAG 3656

Db 2701 GACTGTTATACCTTTGAATCCCTGTGGAATCTACATTTGAAGACAAACATCCTCTGAAG 2760

Qy 3657 TGTATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAAACTGTAAAAATAGCAA 3716

Db 2761 TGTATCAGCAAAAAAATTCAGT - AGATTATCTTTAAAGAAAACTGTAAAAATAGCAA 2819

Qy 3717 CCACTTATGGCACTGTATATATTGTAGACTGTTTTCTCTGTTTTATGCTCTTGTGTAAT 3776

Db 2820 CCACTTATGGCACTGTATATATTGTAGACTGTTTTCTCTGTTTTATGCTCTTGTGTAAT 2879

Qy 3777 CTACTTGACATCATTTTACTCTTGGAAATAGTGGGTGATAGCAAGTATATTCTAAAAAAC 3836

Db 2880 CTACTTGACATCATTTTACTCTTGGAAATAGTGGGTGATAGCAAGTATATTCTAAAAAAC 2939

Qy 3837 TTTGTAAATAAAGTTTTTGTGGCTAAAAATGA 3866

Db 2940 TTTGTAAATAAAGTTTTTGTGGCTAAAAATGA 2969

RESULT 10	
ABX76453	ABX76453 standard; DNA; 2574 BP.
ID	
XX	ABX76453;
AC	
XX	ABX76453;
DT	02-APR-2003 (first entry)
XX	
DE	Lung cancer-associated polynucleotide #317.
XX	
KW	Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW	antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW	chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW	interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX	
OS	Unidentified.
XX	
PN	WO200286443-A2.
XX	
PD	31-OCT-2002.
XX	
PF	18-APR-2002; 2002WO-US012476.
XX	
PR	18-APR-2001; 2001US-0284770P.
PR	10-MAY-2001; 2001US-0290492P.
PR	09-NOV-2001; 2001US-0339245P.
PR	13-NOV-2001; 2001US-0350666P.
PR	29-NOV-2001; 2001US-0334370P.
PR	12-APR-2002; 2002US-0372246P.
XX	
PA	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	
PI	Aziz N, Murray R;
XX	
DR	WPI; 2003-093161/08.
DR	P-PSDB; ABUS6724.
XX	
PT	Detecting a lung cancer-associated transcript in a cell from a patient
PT	for treating lung cancer, by contacting a biological sample from the
PT	patient with a polynucleotide that exhibits increased or decreased
PT	expression in lung cancer.
XX	
PS	Claim 22; Page 433-434; 453pp; English.
XX	
CC	The invention relates to a method for detecting a lung cancer-associated
CC	transcript in a cell from a patient, comprising contacting a biological
CC	sample from the patient with a polynucleotide that selectively hybridises
CC	to a sequence that is at least 80 % identical to a gene that exhibits
CC	increased or decreased expression in lung cancer samples. Lung cancer-
CC	associated polynucleotides and polypeptides are used for identifying a
CC	compound that modulates a lung cancer-associated polypeptide, for
CC	inhibiting proliferation of a lung cancer-associated cell to treat lung
CC	cancer in a patient and for treating a mammal having lung cancer by
CC	administering a modulatory compound identified. The methods are useful
CC	for treating lung cancer, such as small cell lung cancer, non-small cell
CC	lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC	emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC	hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC	bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC	for diagnostic purposes and as targets for screening for therapeutic
CC	compounds that modulate lung cancer, such as antibodies. Sequences
CC	ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC	invention
XX	
SQ	Sequence 2574 BP; 926 A; 463 C; 488 G; 697 T; 0 U; 0 Other;
	Query Match 66.5%; Score 2569.2; DB 7; Length 2574;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 2571; Conservative 0; Mismatches 3; Indels 0; Gaps 0
QY	978 ATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAAATTGATTCCTAATGAACAAGTG 103

Db 1081 GAATCAAGTCTTCTAGCTAAATTAGAAGAACTAAAGAGTATCAAGAACAGAGGTTCCA 1140
QY 2118 GAGAGTAACAGAAACAGTGGCAAGCTTACAGAAAGTCAGAGTGTATTAAACCAGATCCT 2177
Db 1141 GAGAGTAACAGAAACAGTGGCAATCTAAGAGAAAGTCAGAGTGTATTAAACCAGATCCT 1200
QY 2178 GCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAAAAGTTAATACAGAGCAG 2237
Db 1201 GCTGCATCTTCAAAATCACTGGCAGATTCCGGAGTTAGCCCGAAAAGTTAATACAGAGCAG 1260
QY 2238 AAACATACCACTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCACCAACCAATATCA 2297
Db 1261 AAACATACCACTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCACCAACCAATATCA 1320
QY 2298 ACATCTAAATGGTTTGACCCCAAAATCTATTGTAAAGACACCAAGCAGCAATACCTTGGAT 2357
Db 1321 ACATCTAAATGGTTTGACCCCAAAATCTATTGTAAAGACACCAAGCAGCAATACCTTGGAT 1380
QY 2358 GATTACATGAGCTGTTTGTAGAACTCCAGTGTAAAGAAATGACTTTCACCTGCTTGTCTAG 2417
Db 1381 GATTACATGAGCTGTTTGTAGAACTCCAGTGTAAAGAAATGACTTTCACCTGCTTGTCTAG 1440
QY 2418 TTGTCAACACCTTATGGCCCAACCTGCTGTTTCCAGCAGCAACAGCATCAAAATACTTGCC 2477
Db 1441 TTGTCAACACCTTATGGCCCAACCTGCTGTTTCCAGCAGCAACAGCATCAAAATACTTGCC 1500
QY 2478 ACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCCAGCAATGAATGCAITTCGGTT 2537
Db 1501 ACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTCCAGCAATGAATGCAITTCGGTT 1560
QY 2538 AAAGGAAGAAATTTATTCATATTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGTATTT 2597
Db 1561 AAAGGAAGAAATTTATTCATATTAAAGCAGATAGGAAGTGGAGTTCAAGCAAGTATTT 1620
QY 2598 CAGGTGTTAAATGAAGAAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGACGA 2657
Db 1621 CAGGTGTTAAATGAAGAAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGAAGACGA 1680
QY 2658 GATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATTGAATAAACTACAACAA 2717
Db 1681 GATAACCAAACTCTTGATAGTTACCGGAACGAAATAGCTTATTGAATAAACTACAACAA 1740
QY 2718 CACAGTGATAAGATCATCCGACTTTTATGATTATGAATATCAAGGACCAAGTACATCTACATG 2777
Db 1741 CACAGTGATAAGATCATCCGACTTTTATGATTATGAATATCAAGGACCAAGTACATCTACATG 1800
QY 2778 GTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAAGAAATAATCCATTGAT 2837
Db 1801 GTAATGGAGTGTGGAATATTGATCTTAATAGTTGGCTTAAAGAAATAATCCATTGAT 1860
QY 2838 CCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGAGTTTCAACAATCCATCAA 2897
Db 1861 CCATGGGAACGCAAGAGTTACTGGAAAAATATGTTAGAGGAGTTTCAACAATCCATCAA 1920
QY 2898 CATGGCATTTGTTACAGTGATCTTAAACCAAGCTAACTTTCTGATAGTTGATGGAATGCTA 2957
Db 1921 CATGGCATTTGTTACAGTGATCTTAAACCAAGCTAACTTTCTGATAGTTGATGGAATGCTA 1980
QY 2958 AAGCTAATTGATTTGGGATTGCAAAACCAAAATGCAACCAAGATCAACAAAGTGTGTTAAA 3017
Db 1981 AAGCTAATTGATTTGGGATTGCAAAACCAAAATGCAACCAAGATCAACAAAGTGTGTTAAA 2040
QY 3018 GATTCTCAGTTGGCAGAGTTAATTATATGTCACCAAGCAATCAAAAGATATGCTTCTCC 3077
Db 2041 GATTCTCAGTTGGCAGAGTTAATTATATGTCACCAAGCAATCAAAAGATATGCTTCTCC 2100
QY 3078 TCCAGAGAGAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAAGTATGTTTGGTCTCTTA 3137
Db 2101 TCCAGAGAGAATGGGAAATCTAAGTCAAAAGATAAGCCCAAAAGTATGTTTGGTCTCTTA 2160
QY 3138 GGATGTATTTGTACTATATGACTTACGGGAAACACCAATTCAGCAGATAATTAATCAG 3197
Db 2161 GGATGTATTTGTACTATATGACTTACGGGAAACACCAATTCAGCAGATAATTAATCAG 2220

QY 3198 ATTTCTAAATTACATGTCATAATTGATCCTAATCATGAAATGTAATTTCCCGATATTCCA 3257
Db 2221 ATTTCTAAATTACATGTCATAATTGATCCTAATCATGAAATGTAATTTCCCGATATTCCA 2280
QY 3258 GAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAAACAGAGGATA 3317
Db 2281 GAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAAACAGAGGATA 2340
QY 3318 TCCATTCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAACTCATCCAGTTAACCAA 3377
Db 2341 TCCATTCTGAGCTCCTGGCTCATCCCTATGTTCAAATTCAAACTCATCCAGTTAACCAA 2400
QY 3378 ATGCCAAGGGAACCACTGAAGAAATGAATAATGTTCTGGGCCAACTTGTGTTCTGAAT 3437
Db 2401 ATGCCAAGGGAACCACTGAAGAAATGAATAATGTTCTGGGCCAACTTGTGTTCTGAAT 2460
QY 3438 TCTCCTAACTCCATTTTGAAGCTGCTAAACTTTTATATGAACACTATAGTGGTGTGAA 3497
Db 2461 TCTCCTAACTCCATTTTGAAGCTGCTAAACTTTTATATGAACACTATAGTGGTGTGAA 2520
QY 3498 AGTCATAATTCTTCTCATCTCTCAAGACTTTTGAAGAAAGGGGAAAAAATGA 3551
Db 2521 AGTCATAATTCTTCTCATCTCTCAAGACTTTTGAAGAAAGGGGAAAAAATGA 2574

RESULT 11
ADB80469
ID ADB80469 standard; DNA; 2574 BP.

XX ADB80469;

XX 04-DEC-2003 (first entry)

DE Ovarian cancer-associated transcript #17.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection; ds; gene.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 1. .2574
FT CDS /*tag= a

XX WO2002102235-A2.

PN 27-DEC-2002.

PD 18-JUN-2002; 2002WO-US019297.

PF 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA Mack DH, Gish KC;
PI WPI; 2003-167431/16.
XX P-PSDB; ADB80470.

DR Detecting an ovarian cancer-associated transcript in a cell from a
XX patient, comprises contacting a biological sample from the patient with a
XX polynucleotide that hybridizes to an ovarian cancer gene.

PS Claim 10; Page 286-287; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a

CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selection mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.
XX
SQ Sequence 2574 BP; 926 A; 463 C; 488 G; 697 T; 0 U; 0 Other;

Query Match 66.5%; Score 2569.2; DB 9; Length 2574;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2571; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 978 ATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAATTGATTCCTAATAATGAACAAAGTG 1037
Db 1 ATGGAATCCGAGGATTTAAGTGGCAGAGAATTGACAATTGATTCCTAATAATGAACAAAGTG 60

QY 1038 AGAGACATTAATAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAA 1097
Db 61 AGAGACATTAATAAATAAGTTTAAAAATGAAGACCTTACTGATGAACCTAAGCTTGAATAAA 120

QY 1098 ATTTCTGCTGATACACAGATAAATCGGGAACCTGTAACCAAAATTAATGATGATGGCAAA 1157
Db 121 ATTTCTGCTGATACACAGATAAATCGGGAACCTGTAACCAAAATTAATGATGATGGCAAA 180

QY 1158 AACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAAAACAGTGTTCGGCTAAGT 1217
Db 181 AACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAAAACAGTGTTCGGCTAAGT 240

QY 1218 GATGCTCTTTTAAATAAATTAATGATGGTTCGTTACAGTCAAGCAATTGAAGCGCTTCCCCCA 1277
Db 241 GATGCTCTTTTAAATAAATTAATGATGGTTCGTTACAGTCAAGCAATTGAAGCGCTTCCCCCA 300

QY 1278 GATAAATATGCCAAATAGAGATTTTGTCTAGATTAATCAAGTGAGATTGCTGAATTAATA 1337
Db 301 GATAAATATGCCAAATAGAGATTTTGTCTAGATTAATCAAGTGAGATTGCTGAATTAATA 360

QY 1338 GCTATTCAAGAGCCAGATGATGACGTCAGTACTTTTCAAAATGGCCAGCAAACTGCAAG 1397
Db 361 GCTATTCAAGAGCCAGATGATGACGTCAGTACTTTTCAAAATGGCCAGCAAACTGCAAG 420

QY 1398 AAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACCTGTACAGAGTAATGTCAA 1457
Db 421 AAATTTGCTTTTGTTCATATATCTTTTGCACAAATTTGAACCTGTACAGAGTAATGTCAA 480

QY 1458 AAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTTGAGCAGTACCACAGAAATGCTG 1517
Db 481 AAAAGTAAACAACTTCTTCAAAAAGCTGTAGAACGTTGAGCAGTACCACAGAAATGCTG 540

QY 1518 GAAATTCGCTCGGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGGAGGAAAG 1577
Db 541 GAAATTCGCTCGGGAATTTAAACCTCCAAAAAAGCAGCTGCTTTTCAGAGGAGGAAAG 600

QY 1578 AAGAAATTTATCAGCATCTACGGTATTAACCTGCCAAGAATCATTTTCGGTTTCACTTGGG 1637
Db 601 AAGAAATTTATCAGCATCTACGGTATTAACCTGCCAAGAATCATTTTCGGTTTCACTTGGG 660

QY 1638 CATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT 1697
Db 661 CATTTACAGAAATAGGAACAACAGTTGTGATTCAGAGGACAGACTACTAAAGCCAGGTTT 720

QY 1698 TTATATGGAGAGAACATGCCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGA 1757
Db 721 TTATATGGAGAGAACATGCCACCAAGATGCAGAAATAGGTTACCGGAATTCATTGAGA 780

QY 1758 CAACTAAACAAAACCTAAACAGTCAATGCCCCATTTGGAAGAGTCCCAGTTAACCTTCTAAAT 1817
Db 1817 CAACTAAACAAAACCTAAACAGTCAATGCCCCATTTGGAAGAGTCCCAGTTAACCTTCTAAAT 1817

Db 781 CAAACTAACAAAACTAAACAGTCATGCCCATTTGGAAGAGTCCCAGTTAACCTTCTTAAT 840
QY 1818 AGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGAAAAGACAA 1877
Db 841 AGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTATGAAAAGACAA 900
QY 1878 ACCTCTAGATCAGAAATGCCAGATTTTGGTTGCTGCTGGATCTAAACCAAGTGGAATGAT 1937
Db 901 ACCTCTAGATCAGAAATGCCAGATTTTGGTTGCTGCTGGATCTAAACCAAGTGGAATGAT 960
QY 1938 TCCTGTGAATTAAGAAATTAAGTCTGTTTCAAAATAGTCAATTTCAAGGAACCTCTGGTG 1997
Db 961 TCCTGTGAATTAAGAAATTAAGTCTGTTTCAAAATAGTCAATTTCAAGGAACCTCTGGTG 1020
QY 1998 TCAGATGAAAAAGAGTTCTGAACCTTATTATTACTGATTCAATAACCTGAAAGATAAAACG 2057
Db 1021 TCAGATGAAAAAGAGTTCTGAACCTTATTATTACTGATTCAATAACCTGAAAGATAAAACG 1080
QY 2058 GAATCAAGTCTTCTAGCTAAATTAAGAAATTAAGAGAACTAAAGAGTATCAAGAACAGAGTTCCA 2117
Db 1081 GAATCAAGTCTTCTAGCTAAATTAAGAAATTAAGAGAACTAAAGAGTATCAAGAACAGAGTTCCA 1140
QY 2118 GAGAGTAACCCAGAAACAGTGGCAAGCTAAGAGAAAGTCAAGAGTGTATTAACAGAAATCCT 2177
Db 1141 GAGAGTAACCCAGAAACAGTGGCAATCTAAGAGAAAGTCAAGAGTGTATTAACAGAAATCCT 1200
QY 2178 GCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCCGAAAGTAAATACAGAGCAG 2237
Db 1201 GCTGCATCTTCAAATCACTGGCAGATTCGGGAGTTAGCCGAAAGTAAATACAGAGCAG 1260
QY 2238 AAACATACCACTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCAACCAATATCA 2297
Db 1261 AAACATACCACTTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCAACCAATATCA 1320
QY 2298 ACATCTAAATGGTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCAACCAATATCA 2357
Db 1321 ACATCTAAATGGTTTGAGCAACCTGTCTTTTTCAGTTTCAAAACAGTCAACCAATATCA 1380
QY 2358 GATTACATGAGCTGTTTGTAGAACTCCAGTTTGTAAAGAAATGACTTTTCCACCTGCTGTG 2417
Db 1381 GATTACATGAGCTGTTTGTAGAACTCCAGTTTGTAAAGAAATGACTTTTCCACCTGCTGTG 1440
QY 2418 TTGTCAACACCTTATGGCCAACTGCTGCTGTTTCCAGCAGCAACAGCATCAATACTTGGC 2477
Db 1441 TTGTCAACACCTTATGGCCAACTGCTGCTGTTTCCAGCAGCAACAGCATCAATACTTGGC 1500
QY 2478 ACTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAAAATGAATGCATTTCCG 2537
Db 1501 ACTCCACTTCAAAATTTACAGGTTTTTAGCATCTTCTTCAGCAAAATGAATGCATTTCCG 1560
QY 2538 AAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGATTTT 2597
Db 1561 AAAGGAAGAAATTTATTCATATTAAGCAGATAGGAAGTGGAGGTTCAAGCAAGGATTTT 1620
QY 2598 CAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAATATGTGAATTAAGAAAGCA 2657
Db 1621 CAGGTGTTAAATGAAAAGAAACAGATATATGCTATAAAATATGTGAATTAAGAAAGCA 1680
QY 2658 GATAACCAAACTCTGTAGATTACCGGAACGAAATAGCTTATTTGAATTAAGAAAGCA 2717
Db 1681 GATAACCAAACTCTGTAGATTACCGGAACGAAATAGCTTATTTGAATTAAGAAAGCA 1740
QY 2718 CACAGTGAAGATCATCCGACTTTTATGATTATGAATACCGGACAGTACATCTACATG 2777
Db 1741 CACAGTGAAGATCATCCGACTTTTATGATTATGAATACCGGACAGTACATCTACATG 1800
QY 2778 GTAAATGGAGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAAGAAAGAAATCCATGAT 2837
Db 1801 GTAAATGGAGTGTGGAATATTTGATCTTAAATAGTTGGCTTAAAGAAAGAAATCCATGAT 1860
QY 2838 CCATGGGAACGCAAGAGTTACTGGAAATATGTTAGAGGCGAGTTTACACAAATCCATCAA 2897
Db 1861 CCATGGGAACGCAAGAGTTACTGGAAATATGTTAGAGGCGAGTTTACACAAATCCATCAA 1920

QY 2898 CATGGCATTGTTCCAGTGATCTTTAAACCAGCTAACTTTCGATAGTTGATGGATGCTA 2957
Db |||||
QY 1921 CATGGCATTGTTCCAGTGATCTTTAAACCAGCTAACTTTCGATAGTTGATGGATGCTA 1980
Db |||||
QY 2958 AAGCTAATTGATTTGGGATTGCAAAACCAAAATGCAACCAGATACAAAGTGTGTTAAA 3017
Db |||||
QY 1981 AAGCTAATTGATTTGGGATTGCAAAACCAAAATGCAACCAGATACAAAGTGTGTTAAA 2040
QY 3018 GATTCTCAGTTGGCACAGTTAAATTATATGCCACCAGAAAGCAATCAAGATATGCTTCC 3077
Db |||||
QY 2041 GATTCTCAGTTGGCACAGTTAAATTATATGCCACCAGAAAGCAATCAAGATATGCTTCC 2100
QY 3078 TCCAGAGAGAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAAGTGATGTTTGGTCTTA 3137
Db |||||
QY 2101 TCCAGAGAGAATGGGAAATCTAAGTCAAAAGATAAGCCCCAAAAGTGATGTTTGGTCTTA 2160
QY 3138 GGATGTATTTTGTACTATATGACTTACGGGAAAACACCATTTTCAGCAGATAATTATCAG 3197
Db |||||
QY 2161 GGATGTATTTTGTACTATATGACTTACGGGAAAACACCATTTTCAGCAGATAATTATCAG 2220
QY 3198 ATTTCTAAATTACATGCCATAATTGATCCTTAATCATGAATTGAATTTCCCGATATTCCA 3257
Db |||||
QY 2221 ATTTCTAAATTACATGCCATAATTGATCCTTAATCATGAATTGAATTTCCCGATATTCCA 2280
QY 3258 GAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAACACAGGATA 3317
Db |||||
QY 2281 GAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCAAACACAGGATA 2340
QY 3318 TCCATTCTGAGTCTGCTCATCCATATGTTTCAAATTCAAACTCATCCAGTTAACCAA 3377
Db |||||
QY 2341 TCCATTCTGAGTCTGCTCATCCATATGTTTCAAATTCAAACTCATCCAGTTAACCAA 2400
QY 3378 ATGGCCAAGGAAACCACTGAAGAAATGAAATATGTTCTGGCCCAACTTGTGTTCTGAAT 3437
Db |||||
QY 2401 ATGGCCAAGGAAACCACTGAAGAAATGAAATATGTTCTGGCCCAACTTGTGTTCTGAAT 2460
QY 3438 TCTCTAACTCCATTTGAAAGCTGCTFAAAACTTTTATATGAACACTATAGTGGTGA 3497
Db |||||
QY 2461 TCTCTAACTCCATTTGAAAGCTGCTFAAAACTTTTATATGAACACTATAGTGGTGA 2520
QY 3498 AGTCATAATCTTCTCATCTCCAAGACTTTTGAAAAAAAAGGGGAAAAAATGA 3551
Db |||||
2521 AGTCATAATCTTCTCATCTCCAAGACTTTTGAAAAAAAAGGGGAAAAAATGA 2574

RESULT 12
ADC37168
ID ADC37168 standard; DNA; 2860 BP.
XX
AC ADC37168;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 1.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; murine; gene; ds.
XX
OS Mus musculus.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASahi KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR P-PSDB; ADC37169.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 1; 938pp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 2860 BP; 914 A; 601 C; 628 G; 716 T; 0 U; 1 Other;
Query Match 42.2%; Score 1633; DB 9; Length 2860;
Best Local Similarity 76.8%; Pred. No. 0;
Matches 2192; Conservative 0; Mismatches 560; Indels 103; Gaps 12;
QY 971 TGTAGAAATGGAATCCGAGGATTTAAGTGGCAGAGAAATTGACAAATTGATTCCATAATGAA 1030
Db |||||
49 TTTAGAAATGGAGGCTGAAGAGTTAATTGGCAGCAGTGTGACGATTGATTCCATCATGAG 108
QY 1031 CAAAGTGAGAGACATTAATAAATAAGTTTAAAAAATGAAGACCTTACTGATGAACCTAAGCTT 1090
Db |||||
109 CAAAATGAGAGATATTAATAAATAAG---ATAAATGAAGACTGTACTGATGAGCTAAGCTT 165
QY 1091 GAATAAAATTTCTGTGATACTACAGATAAATACTGGGAACCTGTTAACCAAAATTATGATGAT 1150
Db |||||
166 GTCTAAATCTGTG-----CCGATCACACCCGAAACTGTTAACCAAAATTATGAGGCT 216
QY 1151 GGCAAAACAACCCAGAGGACTGGTTGAGTTTGTCTCAAACTAGAGAAAAACAGTGTTC 1210
Db |||||
217 TGGGAACACCCCAAGAGAACTGGTTGAATTTCTTGTGAACTAGAGAAAAACAGCTCACC 276
QY 1211 GCTAAGTGATGCTCTTTTAAATAAATTAATGATTGGTCTGTTACAGTCAAGCAATGAAGCGCT 1270
Db |||||
277 TCTAATGACGATCTTTTAAATAAGCTGATTGGTGGTATAGTCAAGCGATTGAAGTACT 336
QY 1271 TCCCCCAGATAAATATGGCCAAAATGAGAGTTTTGTCTAGAAATTCAGTGAATTTGCTGA 1330
Db |||||
337 TCCTCCAGATAAATACGGCCAGAAATGAGAGCTTTGCTCGAATACAAAGTGAAGACTTGTCTGA 396
QY 1331 ATTAAAGCTATTCAAGAGCCAGATGATGCACGTGACTACTTTCAAATGGCCAGAGCAA 1390
Db |||||
397 ACTAAAGCTATTCAAGAGCCTGATGATGCCCGTGACTACTTCCAGATGGCCAGGAA 456
QY 1391 CTGCAAGAAATTTGCTTTTTCATATATCTTTTGCACAAATTGAACTGTCAACAGGTAA 1450
Db |||||
457 CTGCAAGAAAGTTTGTCTTTTGTGCACGTATCTTTTGCACAGTTGAACTGTCTCAAGGCAA 516
QY 1451 TGTCAAAAAAGTAAACAACTTCTTCAAAAAAGCTGTAGAACGTGGAGCAGTACCCTAGA 1510
Db |||||
517 TCTTAAAAAAGTGAGCAGCTTCTTCATAAAGCTGTAGAGACTGGGCGGTGCCGTGCA 576
QY 1511 AATGCTGGAATTTGCCCTCGCGAATTTAAACCTCCAAAAAAGCAGCTGCTTTCAGAGGA 1570
Db |||||
577 GATGCTGGAGACGCCCATGCGTAACTTACACCTCCAGAAAAAGCAGCTGCTTCCCGAGGA 636
QY 1571 GGAAAAAGAGAAATTATCAGCATCTACGGTATTAACTGCCCAAGAATCATTTTCCGTTTC 1630
Db |||||
637 GGACAAGAGAGTGTGTGAGCATCGACAGTACTAAGTGTCCCAAGAGCCGTTTCTCCAGCTC 696

QY 1631 ACTTGGGCATTTACAGAAATAGGAACAACACAGTTGTGATTCACAGAGGACAGACTACTAAAGC 1690
|||||
Db 697 ACTTGGAAATGTACAGAAATAGGAGCATCAGCTGTGAGTCCAGAGGACAGGCTGGGGCAGC 756
QY 1691 CAGGTTTTATATGGAGAGAACTGCCACCAAGATGCGAGAAATAGGTTACCGGAATTC 1750
|||||
Db 757 CAGGGTTTTATATGGAGAGAACTTGCCTCCACAAGATGCCGAAGTGAAGGCATCAAAACCC 816
QY 1751 ATTGAGACAAACTAACAACAACTAAACAGTCAATGCCCATTTGGGAAGAGTCCCAGTTAACTT 1810
|||||
Db 817 CTTCAAGCAGACTCAGCGAGCTAAACGGTCAATGCCCTTTGGGAAGAGTCCCAGTCAATCT 876
QY 1811 TCTAAATAGCCAGATTTGTGATGTGAAGACAGATGATTCAGTTGTACCTTGTGTTATGAA 1870
|||||
Db 877 TCTAAACAGCCAGATTTCTATGTGAAGACAGATAGCTCAGCTGTGACACAGTTAAACAA 936
QY 1871 AAGACAAACCTCTAGATCAGAAATGCCGAGATTTGGTTGTGCTGGATCTAAACCAAGTGG 1930
|||||
Db 937 AAGAAAGGGGT CAGGACCCAGA---CCGAGACGGGATTTGCCCGGCTCCAGACCACGTTGG 993
QY 1931 AAATGATCTCTGTGAATTAAGAAATTTAAAGTCTGTTCAAAATAGTCAATTTCAAGGAACC 1990
|||||
Db 994 CAGTGATCTCTATGAACCTGAGAGGTTTAAAGCCCATTAACCTATCTATTTGAAAGACTC 1053
QY 1991 TCTGGTGCAGATGAAAAGAGTCTTGAACCTTATTAATTAATTAATTAATTAATTAATTAATTA 2050
|||||
Db 1054 TTTGGTGTCCAATGAAAAGAGTCTGAAC---TTATGTCTGATTTAATAGCCTTGAAGAG 1110
QY 2051 TAAACCGGAATCAAGTCTTCTAGCTAAATTAGAAGAAACTAAAGAGTATCAAGAACCCAGA 2110
|||||
Db 1111 TAAACACAGATTCAAG---TCTAACAATAATTGGAAGAAACTAAG-----CCAGA 1155
QY 2111 GGTTCACAGAGTAACCCAGAAACAGTGGCAAGCTAAAGAGAAAGTCAGAGTGTATTAACCA 2170
|||||
Db 1156 GATTGAGAAAGAGGCCCTGCGAGTGGCAGTCTACCAGAAAGCCCGAGTGTGTTCCA 1215
QY 2171 GAATCCTGCTGATCTTCAATCACTGGCAGATTCGGAGTTAGCCCGAAAAGTTAATAC 2230
|||||
Db 1216 GAACCTCTGCTGCTTTGACCCCTCGGCACTGTCAGATGTCAACCCGGAAGCTGACAA 1275
QY 2231 AGAGCAGAAACATACCACCTTTTGAGCAACCTGCTCTTTTCAGTTTCAAAACAGTCACACC 2290
|||||
Db 1276 AG-----AGTCAACACC 1287
QY 2291 AATATCAACATCTAAATGGTTTGAACCCAAATCTATTGTGAAGACACCAAGCAGCAATAC 2350
|||||
Db 1288 AATATCAGTCTCTAAATGGCTTGATCCAAAGTCTGCTTGTGAGACACCTAGTAGCAGCTC 1347
QY 2351 CTTGGATGATTACATGAGCTGTTTTAGAACTCCAGTTGTAAAGAAATGACTTTCCACCTGC 2410
|||||
Db 1348 CTTGGATGATTACATGAAATGTTTTAAAGACTCCAGTTGTAAAGAAATGACTTTCCACCTGC 1407
QY 2411 TTGTCAGTTGTCAACACCTTATGGCCAACTGCTGTTTTCCAGCAGCAACAGCATCAAT 2470
|||||
Db 1408 CTGTCCATCATCAACACCTTACAGCCAGCTTGGCCGCTCCAGCAGCAACAGCAGCAGGG 1467
QY 2471 ACTTGCCACTCCACTTCAAAATTTACAGGTTTTAGCATCTTCTTACGCAAAATGAATGCAT 2530
|||||
Db 1468 ACTCAGCACTCCTCTTCAAGCTTGCAAGTTTCAAGTTTCAAGTTTCTTCAATAAATGAATGCAT 1527
QY 2531 TTCGGTTAAAGGAAGAAATTTATTCATATTAAAGCAGATAGGAAGTGGAGGTTTCAAGCAA 2590
|||||
Db 1528 TTCAGTTAAACGGGAAGAAATTTATTCATATTAAAGCAGATAGGCAGTGGAGGTTCCAGTAA 1587
QY 2591 GGTATTTAGGTTGTTAAATGAAAAGAAACAGATATATGCTATAAAATATGTGAACCTTAGA 2650
|||||
Db 1588 GGTGTTTCAGGTATTGAATGAGAAAGAAACAGATAAACGCTATCAAAATATGTGAACCTTAGA 1647
QY 2651 AGAAGCAGATAACCAACTCTTGTAGATGTTACCGGAACGAAATAGCTTATTTGAATAAAT 2710
|||||
Db 1648 AGACGCCGATAGCCAAACTATTGAGAGCTACCGCAACGAGATAGCGTTTTTGAACAAACT 1707
QY 2711 ACAACACACAGTGATAAGATCATCCGACTTTATGATTATGAAATCAGGACCAGTAGTACAT 2770

Db 1708 ACAGCAACACAGTGATAAGATCATCCGCTCTATGATTATGAAATCACCGAGCAGTACAT 1767
QY 2771 CTACATGGTAATGGAGTGTGGAAATATTGATCTTAATAGTTGGCTTAAAAAGAAAAATC 2830
|||||
Db 1768 CTACATGGTAATGGAAATGTGGAAACATTGACCTAAATAGTTGGCTTAAAAAGAAAAATC 1827
QY 2831 CATTGATCCATGGGAACGCAAGAGTTACTGGAATAATATGTTAGAGCAGTTTACACAAT 2890
|||||
Db 1828 CATCAATCCATGGGAACGCAAGAGCTACTGGAATAATCATGTTGGAGCAGTACACATAAT 1887
QY 2891 CCATCAACATGGCATTTGTTCAAGTGTACTTAAACCAGTAACTTTCTGTATAGTTGATGG 2950
|||||
Db 1888 CCATCAGCATGGTATTGTTTCATAGTGAATCTGAAGCTGTAACTTTGTGATAGTTGACGG 1947
QY 2951 AATGCTAAAGCTAATTGATTTTGGGATTTGCAATGCAACCAAAATGCAACCAAGCAATCAAAGATAT 3010
|||||
Db 1948 AATGCTAAAGCTAATTGATTTTGGGATTTGCAATGCAACCAAAATGCAACCAAGCAATCAAAGATAT 2007
QY 3011 TGTTAAAGATTTCTCAGGTTGGCACAGTTAAATATATATATATATATATATATATATATATATAT 3070
|||||
Db 2008 TGTTAAAGATTTCTCAGGTTGGCACAGTTAACTATATATATATATATATATATATATATATAT 2067
QY 3071 GTCTTCTCCAGAGAGAAATGGAATCTAAGTCAAAAGATAAGCCCAAAAGTATGTTTG 3130
|||||
Db 2068 GTCTTCTTCAAGAGAAATTCGAAATCAGGACCAAGGTAAGTCCCAAGAGTATGTTCTG 2127
QY 3131 GTCTTAGGATGTTTCTACTATATATGATTTACGTTACGGGAAACACCAATTTTCCAGATATAT 3190
|||||
Db 2128 GTCTTGGGGTGCAATTTGTACTACATGACITATGGAGGAGCCCATTTTCCAGCATCAT 2187
QY 3191 TAATCAGATTTCTAAATTACATGCCATAATTTGATCCTAATCATGAATTTGAATTTCCCGA 3250
|||||
Db 2188 CAATCAGGCTCTAAACCTGACGCCATAATCAACCTGCTCATGAGTTGAATTTCCCGA 2247
QY 3251 TATTCAGAGAAAGATCTTCAAGATGTGTTAAAGTGTGTTTAAAGGGACCCCAAAACA 3310
|||||
Db 2248 GATTCGGAAAAAGATCTTCGAGACGTGTTAAAGTGTGTTTAAAGGAGAACCCCTAAAGA 2307
QY 3311 GAGGATATCCATTCCTGAGCTCCTGGCTCATCCATATGTTCAAATTCAAACTCATCCAGT 3370
|||||
Db 2308 GAGGATATCTATCCCTGAGCTTCTCACACATCCGTATGTTCAAATTCAGCCCCATCCAGG 2367
QY 3371 TAACCAAAATGGCCAAAGGAACCACTGAAGAAATGAAATATGTTCTGGGCCAACTTGTGG 3430
|||||
Db 2368 CAGCCAAATGGCTAGGGGAGCCACTGATGAATGAAATATGTTGGGTCAACTTGTGG 2427
QY 3431 TCTGAATCTCTAACTCCATTTTGAAGCTGCTAAAACTTTATATGAACACTATAGTGG 3490
|||||
Db 2428 TCTGAATCTCTAACTCCATCTTGAAGCTGCTAAAACTTTGATGAACAGTTATAATG 2487
QY 3491 TGGTGAAGTCTATAATCTTCTCATCTCCAGACTTTTGAAGAAATAAGGGGAAAAAAATG 3550
|||||
Db 2488 TGGTGAAGTCAAGATTTCTGCTCATCCAGACTTTTGACAAAAAGAGAGAAAGAAAGT 2547
QY 3551 ATTTGCAATTATTCGTAATGTCAGATAGGAGTATAAAATATATTTGACTGTTATACTCT 3610
|||||
Db 2548 ATGCACAGCTACGTACAAACCAAGA-----ACACTAGATTGTTTCTCTGCCATACTCT 2601
QY 3611 TGAATCCCTGTGGAATCTA-CATTTGAAGACACATCACTCTGAAGTGT----- 3660
Db 2602 TGAATCTGTGAGAAATCTACAGTTGGAAACCAACCTCACCTGGATTTTATCAGTTAAAA 2661
QY 3661 --ATCAGCAAAAAAATTCAGTGAGATTATCTTTAAAGAAACCTGTAATAATAGCAACC 3718
Db 2662 AAACAAACAAACAACTTCAGTAGATTATCTCTCAAAAGGAAGCTGTAAAGTT---AACC 2718
QY 3719 ACTTATGGCACTGTATATATTTAGACTTGTGTTTCT-CTGTTTTATGCTCTTGTGTAATC 3777
Db 2719 ACTCATAGCACTGTGTATATTAATTTATAGAGTGTGCTTTTCTTTTATGCTTTTCTGTA 2778
QY 3778 TACTTGACATCATTTTACTCTTGGAAATAGTGGGTG 3812

Db 2779 AATCTGCTAATGTTTACGTTTAGAACAGTGAATG 2813

RESULT 13

ADA02894/c

ID ADA02894 standard; DNA; 32404 BP.

XX AC ADA02894;

XX DT 06-NOV-2003 (first entry)

XX DE Human BLR1 carcinoma associated gene, SEQ ID NO:1412.

XX KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast; prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening; gene; ds.

XX OS Homo sapiens.

XX PN WO2003057146-A2.

XX PD 17-JUL-2003.

XX PF 26-DEC-2002; 2002WO-US041414.

XX PR 26-DEC-2001; 2001US-00035832.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX DR WPI; 2003-587068/55.

PT New recombinant nucleic acid encoding carcinoma associated protein, useful for preparing compositions for treating carcinomas.

PS Claim 1; SEQ ID NO 1412; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 32404 BP; 7711 A; 8273 C; 9097 G; 7303 T; 0 U; 20 Other;

Query Match 10.2%; Score 395.6; DB 8; Length 32404;

Best Local Similarity 81.3%; Pred. No. 4e-83;

Matches 499; Conservative 0; Mismatches 104; Indels 11; Gaps 3;

QY 9 TTTTCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGCATGGCA 68

Db TTTTCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGCATGGCA 31915

QY 69 CAATCTCAGCTTACTGCAACCTCCGCCCTCCGGGTTCAAGCGATCTCCTGCCTCAGCCT 128

TTTCTTTTGTGAGATGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGCATGGCA 31915

Db 31914 CGAACTGGCTTACTGCAACCTCTGCCTCCTGAGTTCAAGCGATTTTCTCCTCAGCCT 31855

QY 129 CTCAAGTAGCTGGGATTACAGGCATGTGCCACACCCCTGGCTAACTAATTTCTTTTCTA 188

Db 31854 CCGAGTAGCTGGGATTACAGGCACCTGCCACACCGCTGGCTAA-----TTTGTGATT 31800

QY 189 TTTAGTAGAGATGGGGTTTCAACCATGTTGGTCAAGGCTGGTCTTGAACCTCCTGACCTCAGG 248

Db 31799 TTTAGCAGAGGTGGAGTTTAGCATGTTGGCCAGACCGGTCTCTGAACCTCCTGACCTCAGG 31740

QY 249 TGATCCACTTGCCTTGGCCTCCCAAAGTCTAGGATTACAGCCGCGAAACTGTGCCTGGC 308

Db 31739 TGATACACCTGCCTCAGCCCCCAAGTCTGGGATTACAGGCGTGAGCC--ACCCGGCC 31682

QY 309 TGATCTTTTGTGTTGGATTGTTTGAACAGGGT---CTCCCTTGGTCCGCCAGGC 364

Db 31681 TGATCTTTTGTGTTGGATTGTTTGAACAGGGTCTCACTCACTCTGTGTCGCCAGGC 31622

QY 365 TGGAGTGCAGTGGTGGATCTTGGCTCACTATAACCTCCACCTCCTGGTTTCAAGTGATC 424

Db 31621 TGGAGTGCAGTGGCAAACTCTCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTAAT 31562

QY 425 CTCCACTTTAGCCTCCTGAGTAGCTGTGATTACAGGCGTGCAACACACCCGGCTAA 484

Db 31561 CTCTGTCTCAGCCTCCCAAAGTAGCTGGGATTACAGGCATCGGTCAACACCGCTGGCTAA 31502

QY 485 TTTTGTATTTTATTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGTTCTCAAACTCC 544

Db 31501 TTTTGTATTTTATTAGAGATGGAGTTTCACTATGTTGGCCAGGCTGTTCTTGAACCTCC 31442

QY 545 TGGACTCAAGGGATCCGCTGCCTCCACTTCCAAAGTCCCAGAGATTACAGGTGTGAGTC 604

Db 31441 TGCCTCAAGTGATCCGCCGCTCGGCTCCGCTCCAAAGTGTAGGATTACAGGCATAAGCC 31382

QY 605 ACCATGCCTGACCT 618

Db 31381 ACCGTGCCCGGCT 31368

RESULT 14

ADB72632/c

ID ADB72632 standard; DNA; 32404 BP.

XX AC ADB72632;

XX DT 04-DEC-2003 (first entry)

XX DE Human BLR1 gene.

XX KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas; cancer; neoplasm; adenocarcinoma; sarcoma; gene.

XX OS Homo sapiens.

XX PN WO2003008583-A2.

XX PD 30-JAN-2003.

XX PF 26-DEC-2001; 2001WO-US051291.

XX PR 02-MAR-2001; 2001US-00798586.

XX PR 23-OCT-2001; 2001US-00004113.

XX PR 08-NOV-2001; 2001US-00052482.

XX PR 30-NOV-2001; 2001US-00997722.

XX PR 20-DEC-2001; 2001US-00034650.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW, Engelhard EK;

XX DR WPI; 2003-239337/23.

XX PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,

PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 460; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a human gene of the invention.
XX
SQ Sequence 32404 BP; 7711 A; 8273 C; 9097 G; 7303 T; 0 U; 20 Other;

Query Match 10.2%; Score 395.6; DB 9; Length 32404;
Best Local Similarity 81.3%; Pred. No. 4e-83;
Matches 499; Conservative 0; Mismatches 104; Indels 11; Gaps 3;

QY 9 TTTTATTTTATTTTGGAGTGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGAATGGCA 68
Db 31974 TATTTATTTATTTTGAATCGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGAATGGCA 31915

QY 69 CAATCTCAGCTTACTGCAACCTCCGCTCCCGGTTCAAGCGATTCTCCTGCCTCAGCCT 128
Db 31914 CGAACTCGGCTTACTGCAACCTCTGCCTCTGAGTTCAAGCGATTCTCCTGCCTCAGCCT 31855

QY 129 CTCAAGTAGCTGGGATTACAGGCATGTGCCACCATCCCGGTTCAAGCGATTCTCCTGCCTCAGCCT 188
Db 31854 CCCGAGTAGCTGGGATTACAGGCATGTGCCACCATCCCGGTTCAAGCGATTCTCCTGCCTCAGCCT 31800

QY 189 TTTAGTAGAGTGGGTTTCCACATGTTGGTCAAGGTGGTCTTGAACCTCCTGACCTCAGG 248
Db 31799 TTTAGCAGAGTGGAGTTTATGATGTTGGCCAGGCTGGAGTGAATGGCA 31740

QY 249 TGATCCACTTGCCTTGGCCTCCCAAAGTGTAGGATTACAGCCGTGAAGTGTGCTGGC 308
Db 31739 TGATACACCTGCCTCAGCCCCCAAAAGTGTGGGATTACAGGCGTGAGCC--ACCGGGCC 31682

QY 309 TGATTCCTTTTGTGTTGGATTGTTGAAACAGGGT----CTCCCTTGGTGGCCAGGC 364
Db 31681 TGATTCCTTTTGTGTTGGATTGTTGAAACAGGGT----CTCCCTTGGTGGCCAGGC 31622

QY 365 TGGAGTGCAGTGGTGGCATTCTTGGCTCACTATAACCTCCACCTCTCGGTTTCAAGTGATC 424
Db 31621 TGGAGTGCAGTGGCACAATCTCGGCTCACTGTCAACCTCCACCTCTCGGTTTCAAGTAAT 31562

QY 425 CTCCCACTTTAGCCTCTCTGAGTAGCTGTGATTACAGGCGTGACACACACCCGGCTAA 484
Db 31561 CTCTGCTTCAAGCTCCCAAGTAGCTGGGATTACAGGCGATGCGTCAACACGCTGGCTAA 31502

QY 485 TTTTGTATTTTATTTAGACAGGGTTTACCATGTTGGCCAGGCTGTTCTCAAACTCC 544
Db 31501 TTTTGTATTTTATTTAGTAGAGTGGAGTTTCAACCATATATGGCCAGGCTGTTGAAGTCC 31442

QY 545 TGGACTCAAGGATCCGCTGCTCCACTTCCCAAGTCCCGAGATTACAGGTGTGAGTC 604
Db 31441 TGCCCTCAAGTATCCGCGCGCTCGGCTCCCAAGTGTAGGATTACAGGCGATAAGCC 31382

QY 605 ACCATGCTGACCT 618
Db 31381 ACCGTGCCGCGCT 31368

RESULT 15
ADC85373/c
ID ADC85373 standard; DNA; 32404 BP.
XX
AC ADC85373;
XX
DT 01-JAN-2004 (first entry)
XX
DE Mouse Blr1 coding sequence.
XX

KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.
XX
OS Mus sp.
XX
PN WO2003045230-A2.
XX
PD 05-JUN-2003.
XX
PF 02-DEC-2002; 2002WO-US038582.
XX
PR 30-NOV-2001; 2001US-00997722.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
XX WPI; 2003-513603/48.
XX
CC The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC ADC85514 represent CA genes of the invention.
XX
SQ Sequence 32404 BP; 7711 A; 8273 C; 9097 G; 7303 T; 0 U; 20 Other;

Query Match 10.2%; Score 395.6; DB 9; Length 32404;
Best Local Similarity 81.3%; Pred. No. 4e-83;
Matches 499; Conservative 0; Mismatches 104; Indels 11; Gaps 3;

QY 9 TTTTATTTTATTTTGGAGTGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGAATGGCA 68
Db 31974 TATTTATTTATTTTGAATCGGAGTTTCACTCTTGTGGCCAGGCTGGAGTGAATGGCA 31915

QY 69 CAATCTCAGCTTACTGCAACCTCCGCTCCCGGTTCAAGCGATTCTCCTGCCTCAGCCT 128
Db 31914 CGAACTCGGCTTACTGCAACCTCTGCCTCTGAGTTCAAGCGATTCTCCTGCCTCAGCCT 31855

QY 129 CTCAAGTAGCTGGGATTACAGGCATGTGCCACCATCCCGGTTCAAGCGATTCTTCTTCTA 188
Db 31854 CCCGAGTAGCTGGGATTACAGGCATGTGGCCAGGCTGGAGTGAATGGCA 31800

QY 189 TTTAGTAGAGTGGGTTTCCACATGTTGGTCAAGGTGGTCTTGAACCTCCTGACCTCAGG 248
Db 31799 TTTAGCAGAGTGGAGTTTATGATGTTGGCCAGGCTGGAGTGAATGGCA 31740

QY 249 TGATCCACTTGCCTTGGCCTCCCAAAGTGTAGGATTACAGCCGTGAAGTGTGCTGGC 308
Db 31739 TGATACACCTGCCTCAGCCCCCAAAAGTGTGGGATTACAGGCGTGAGCC--ACCGGGCC 31682

QY 309 TGATTCCTTTTGTGTTGGATTGTTGAAACAGGGT----CTCCCTTGGTGGCCAGGC 364
Db 31681 TGATTCCTTTTGTGTTGGATTGTTGAAACAGGGT----CTCCCTTGGTGGCCAGGC 31622

QY 365 TGGAGTGCAGTGGTGGCATTCTTGGCTCACTATAACCTCCACCTCTCGGTTTCAAGTGATC 424
Db 31621 TGGAGTGCAGTGGCACAATCTCGGCTCACTGTCAACCTCCACCTCTCGGTTTCAAGTAAT 31562

QY 425 CTCCCACTTTAGCCTCTCTGAGTAGCTGTGATTACAGGCGTGACACACACCCGGCTAA 484
Db 31561 CTCTGCTTCAAGCTCCCAAGTAGCTGGGATTACAGGCGATGCGTCAACACGCTGGCTAA 31502

QY 485 TTTTGTATTTTATTTAGACAGGGTTTACCATGTTGGCCAGGCTGTTCTCAAACTCC 544
Db 31501 TTTTGTATTTTATTTAGTAGAGTGGAGTTTCAACCATATATGGCCAGGCTGTTGAAGTCC 31442

QY 545 TGGACTCAAGGATCCGCTGCTCCACTTCCCAAGTCCCGAGATTACAGGTGTGAGTC 604
Db 31441 TGCCCTCAAGTATCCGCGCGCTCGGCTCCCAAGTGTAGGATTACAGGCGATAAGCC 31382

QY 605 ACCATGCTGACCT 618
Db 31381 ACCGTGCCGCGCT 31368

RESULT 15
ADC85373/c
ID ADC85373 standard; DNA; 32404 BP.
XX
AC ADC85373;
XX
DT 01-JAN-2004 (first entry)
XX
DE Mouse Blr1 coding sequence.
XX

Db	31501	TTTTTGTAATTTTAGTAGAGATGGAGTTTCACCAATTATGGCCAGGCTGGTCTTGAACTCC	31442
QY	545	TGGACTCAAGGGATCCGCCTGCCTCCACTTCCCAAAGTCCCGAGATTACAGGTGTGAGTC	604
Db	31441	TGCCCTCAAGTGATCCGCCCGCCTCGGCCTCCCAAAGTTAGGATTACAGGCATAAGCC	31382
QY	605	ACCATGCCCTGACCT	618
Db	31381	ACCGTGCCCCGCCT	31368

Search completed: September 30, 2004, 02:36:07 -
Job time : 1334 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:02:29 ; Search time 134 Seconds
(without alignments)
1773.302 Million cell updates/sec

Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDIKNFKNEDLTDEL.....GGESHNSSSSKTFEKKRGKK 841

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4376	100.0	841	5	ABP54938 Human tyr
2	4376	100.0	841	7	ABU61611 Human Tyr
3	4376	100.0	841	7	ADD49934 Human lun
4	4376	100.0	841	7	ADE40437 Human TTK
5	4376	100.0	841	7	ADE31645 Human 168
6	4376	100.0	857	7	ADE38369 Human pro
7	4373	99.9	857	6	ABU56724 Lung canc
8	4373	99.9	857	7	ADB80470 Ovarian c
9	4373	99.9	857	7	ADC37171 Nuclear f
10	3182.5	72.7	830	7	ADC37169 Nuclear f
11	673.5	15.4	678	5	ABP54948 Schizosac
12	673.5	15.4	678	7	ABU61621 Yeast MPH
13	647.5	14.8	777	3	AAG38500 Arabidops
14	643	14.7	720	3	AAG38501 Arabidops
15	640.5	14.6	638	3	AAG38502 Arabidops
16	637	14.6	630	4	ABB63414 Drosophil
17	603	13.8	735	6	ABJ25382 Aspergill
18	603	13.8	842	6	ABJ25982 Aspergill
19	586.5	13.4	764	5	ABP54947 Saccharom
20	586.5	13.4	764	7	ABU61620 Yeast MPS
21	585.5	13.4	764	2	AAW26727 Yeast Mps
22	558	12.8	690	5	ABP73673 Candida a
23	523.5	12.0	331	3	AAG05720 Arabidops
24	490	11.2	297	3	AAG05721 Arabidops
25	472	10.8	282	3	AAG05722 Arabidops

26	342.5	7.8	749	7	ADC37183	Adc37183 Nuclear f
27	342.5	7.8	809	4	AAM41064	Aam41064 Human pol
28	340	7.8	719	6	ABU07446	Abu07446 Protein d
29	338	7.7	528	5	ABP73734	Abp73734 Candida a
30	338	7.7	717	7	ADC37185	Adc37185 Nuclear f
31	338	7.7	718	4	AAM39278	Aam39278 Human pol
32	336	7.7	1090	3	AAB10946	Aab10946 L. mexica
33	332.5	7.6	250	5	ABB78798	Abb78798 Serine/th
34	332	7.6	598	7	ADC37181	Adc37181 Nuclear f
35	332	7.6	809	6	ABU11737	Abu11737 Human MDD
36	329.5	7.5	749	5	ABU65050	Abu65050 Human NOV
37	328.5	7.5	787	7	ADE38387	Ade38387 Human pro
38	326.5	7.5	748	4	ABG26847	Abg26847 Novel hum
39	325.5	7.4	787	2	AAR76061	Aar76061 Protein k
40	324	7.4	520	4	ABB64044	Abb64044 Drosophil
41	323.5	7.4	588	4	AAB65659	Aab65659 Novel pro
42	319	7.3	231	4	AAG67541	Aag67541 Consensus
43	319	7.3	540	2	AAR76062	Aar76062 Protein k
44	318.5	7.3	431	3	AAB27933	Aab27933 Human sec
45	318.5	7.3	541	3	AAB27995	Aab27995 Human sec

ALIGNMENTS

RESULT 1
ABP54938
ID ABP54938 standard; protein; 841 AA.
XX
AC ABP54938;
XX
DT 13-JAN-2003 (first entry)
XX
DE Human tyrosine threonine kinase.
XX
KW Tyrosine threonine kinase; TTK; protein kinase; enzyme; human;
KW mitotic checkpoint; colon cancer; breast cancer; tumour; cytostatic;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WC200268444-Al.
XX
PD 06-SEP-2002.
XX
PF 21-FEB-2002; 2002WO-US0005278.
XX
PR 21-FEB-2001; 2001US-0271254P:
XX (CHIR) CHIRON CORP.
Reinhard C, Jefferson AB, Chan VW;
WPI; 2002-698650/75.
N-PSDB; ABV73974.

Reducing growth of cancer cells comprises reducing Tyrosine Threonine Kinase (TTK) activity, useful in diagnosing and treating disorders with abnormal expression levels and activity of TTK, such as lung, colon, prostate and ovarian cancer.
Claim 7; Page 71-73; 113pp; English.
The present sequence is the protein sequence of human tyrosine threonine kinase (TTK), a mixed specificity protein kinase that is involved in the cell cycle spindle assembly checkpoint. The invention is based on the finding that TTK is differentially expressed in colon tumour cells relative to healthy tumour colon cells as detected by microarray analysis. Differential expression was confirmed in cell lines derived from various forms of cancer, indicating involvement of TTK in cancer as a more general mechanism. Disruption of TTK function using antisense oligonucleotides decreased proliferation, inhibited anchorage-independent growth, and induced apoptosis of cancer cell lines, including a

CC metastatic breast cancer cell line (MDA-MB-213) and a colorectal
CC carcinoma cell line (SW620). These findings serve as the basis for the
CC materials and methods of the invention for use in diagnosing cancer of a
CC patient, for planning and selection of appropriate therapeutic and/or
CC prophylactic treatment, and for treatment of cancer associated with
CC aberrant TTK levels, e.g. associated with overexpression or
CC overproduction of TTK, by inhibition of gene product production (by
CC decreasing levels of transcription and/or translation), by decreasing TTK
CC activity by decreasing TTK gene product production, and/or by reducing
CC one or more of TTK's kinase activities in cancer cells, especially breast
CC cancer or colon cancer cells. The present invention has the advantage of
CC having the ability to project disease progression based on expression of
CC TTK in a malignant or pre-malignant growth. It allows a more systematic
CC approach for intervention of a cancerous disease

XX Sequence 841 AA;

Query Match 100.0%; Score 4376; DB 5; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.1e-292;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKPKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVFAELKAIQEPDDARDYFQMAR 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 EEEKNLASASTVLTAEQSFSGSLGHLQNRNNSCDRGQTTKARFLYGENMPPQDAEIGYR 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDKVTDSVVPFCFMKQRTSRSECRDLVVPGSKP 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 SGNDSCBLNLKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 PEVPESNQKQWQAKRKSECEINQNPAASSNHQWIPELARKVNTQKHTTPEQVFSVSKQS 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFFPPACQLSTPYGQACFQQQOH 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 QILATPLQNLQVLASSANECISVKGRISYILKQIGSGGSKVFQVLNEKKQIYAIKYVN 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 LEADNOTLDSYRNEIAYLNKLQHQSDKIIRLYDYEITDQYIYVMCEGNIIDNSWLKKK 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDPT 660
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 661 SVVKDSQVGTVNYMPPPAIKDMSSSRENGSKSKISPKSDVMSLGCILYMTYGTPTFQQ 720
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 721 IINQISKLHAIIDPNHEIEFFDPEKLDQVLKCLKRDPKQISIPELLAHPYVQIQTH 780
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 781 PVNQMAKGTTEEMKYVLGVLNPNPNSILKAAKTLVHYSGGESHNSSSSKTFEKKRGK 840
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 841 K 841
Db 841 K 841

RESULT 2

ABU61611

ID ABU61611 standard; protein; 841 AA.

XX AC ABU61611;

XX DT 11-AUG-2003 (first entry)

XX DE Human Tyrosine threonine kinase, TTK.

XX KW Human; enzyme; tyrosine threonine kinase; TTK; cancer; cytostatic;
mitotic checkpoint gene.

XX OS Homo sapiens.

XX PN US2003045491-A1.

XX PD 06-MAR-2003.

XX PF 21-FEB-2002; 2002US-00081119.

XX PR 23-FEB-2001; 2001US-0289813P.

XX PA (REIN/) REINHARD C.

XX PA (JEFF/) JEFFERSON A B.

XX PA (CHAN/) CHAN V W.

PI Reinhard C, Jefferson AB, Chan VW;

XX WPI; 2003-456566/43.

DR N-PSDB; ACA62263.

XX
PT Detecting cancer in a subject, by comparing expression levels of tyrosine
threonine kinase polypeptide or polynucleotide in a subject cell and a
normal cell, where an increase in the expression level in the test cell
is indicative of cancer.

Claim 7; Page 25-27; 79pp; English.

CC The invention relates to detecting cancer (other than ovarian cancer) in
a subject, comprising comparing the expression levels of tyrosine
threonine kinase (TTK, a mitotic checkpoint gene) polypeptide or
polynucleotide in a test cell obtained from the subject and in a normal
non-cancer cell, where an increase in the expression level of TTK protein
or nucleic acid in the test cell compared to that in the normal cell,
indicates the presence of cancer other than ovarian cancer. Also included
are reducing growth of a cancerous cell (by contacting a cancerous cell
with an amount of an agent effective to reduce TTK polypeptide activity
in the cell), an assay for identifying a candidate agent that reduces
growth of a cancerous cell (comprising: (i) detecting the activity of a
TTK polypeptide in the presence of a candidate agent; and (ii) comparing
the activity of TTK polypeptide in the presence of a candidate agent
relative to TTK polypeptide activity in the absence of the candidate
agent), identifying an agent that reduces TTK activity (comprising: (i)
contacting a cancerous cell displaying elevated expression of a TTK-
encoding polynucleotide with a candidate agent; and (ii) determining the
effect of the candidate agent on TTK polypeptide activity) and assessing
the prognosis of a cancerous disease other than ovarian cancer in a
subject (comprising: (i) detecting expression of TTK -encoding
polynucleotide in a test cancer cell of a subject; and (ii) comparing a
level of expression of TTK-encoding polynucleotide in the test cancer
cell with a level of expression of the polynucleotide in a control non-
cancer cell, where the level of expression of TTK in the test cancer cell
is relative to the level of expression in the control non-cancer cell is

CC stimulating an immune response in a patient (comprising administering to
CC the patient the above composition) treating lung cancer in a patient
CC (comprising administering to the patient the above composition and a
CC diagnostic kit (comprising: at least one oligonucleotide cited above; or
CC at least one antibody cited above and a detection reagent, where the
CC detection reagent comprises a reporter group). The composition and
CC methods are useful in diagnosing, preventing and treating cancer,
CC particularly lung cancer. The present sequence is a lung cancer-
CC associated antigen of the invention.
XX
SQ Sequence 841 AA;

Query Match 100.0%; Score 4376; DB 7; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.1e-292;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKPKVEDLTDELSLNKISADTTDSNGTVNQIMMANNPEDWLSLLKLEKNS 60
Db 1 MNKVRDIKNKPKVEDLTDELSLNKISADTTDSNGTVNQIMMANNPEDWLSLLKLEKNS 60
QY 61 VPLSDALLNKLIGRYSOAIEALPPDKYQNESFARIQVFAELKAIQEPDARDYFQMAR 120
Db 61 VPLSDALLNKLIGRYSOAIEALPPDKYQNESFARIQVFAELKAIQEPDARDYFQMAR 120
QY 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180
Db 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180
QY 181 EEEKNLASASTVLTAEQSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
Db 181 EEEKNLASASTVLTAEQSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
QY 241 NSLRQTNKTKQSCFPGRVPVNLNLLSPDCDKVTDSDVVPFCFMKQTSRSECRDLVVPGSKP 300
Db 241 NSLRQTNKTKQSCFPGRVPVNLNLLSPDCDKVTDSDVVPFCFMKQTSRSECRDLVVPGSKP 300
QY 301 SGNDSCBLRLNLSVQNSHFKEPLVSDKSSSELIITDSTILKNKTESILLAKLEETKEYQE 360
Db 301 SGNDSCBLRLNLSVQNSHFKEPLVSDKSSSELIITDSTILKNKTESILLAKLEETKEYQE 360
QY 361 PEVPESNQKQWQAKRKSECCINQNPAASSNHWQIPELARKVNTQKHTTFFQPVFSVSKQS 420
Db 361 PEVPESNQKQWQAKRKSECCINQNPAASSNHWQIPELARKVNTQKHTTFFQPVFSVSKQS 420
QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRPVPVKNDFPPACQLSTPYGQACFQQOQH 480
Db 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRPVPVKNDFPPACQLSTPYGQACFQQOQH 480
QY 481 QILATPLONLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVN 540
Db 481 QILATPLONLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVN 540
QY 541 LEEADNQTLDYSYRNEIAYLNKLQHSDDKIIRLYDYEITDQYIYVMECGNIDLSNWLKKK 600
Db 541 LEEADNQTLDYSYRNEIAYLNKLQHSDDKIIRLYDYEITDQYIYVMECGNIDLSNWLKKK 600
QY 601 KSIDPWKRSYKWNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMPDPT 660
Db 601 KSIDPWKRSYKWNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMPDPT 660
QY 661 SVVKDSQVGTVNYMPPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYGTPTFFQ 720
Db 661 SVVKDSQVGTVNYMPPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYGTPTFFQ 720
QY 721 IINQISKLHAIIDPNHEIEFPDIPKDLQVLCCLKRDPKQKISIPPELLAHYVQIQTH 780
Db 721 IINQISKLHAIIDPNHEIEFPDIPKDLQVLCCLKRDPKQKISIPPELLAHYVQIQTH 780
QY 781 PVNQMAKGTTEEMKYVLGVLGNLSPNSILKAAKTLYEHYSGGESHSSSSSKTFEKKRGK 840
Db 781 PVNQMAKGTTEEMKYVLGVLGNLSPNSILKAAKTLYEHYSGGESHSSSSSKTFEKKRGK 840
QY 841 K 841

Db 841 K 841

RESULT 4

ADE40437

ID ADE40437 standard; protein; 841 AA.

XX AC ADE40437;

XX DT 29-JAN-2004 (first entry)

XX DE Human TTK protein kinase (gene ID 1682) protein.

XX KW AIDS; acquired immunodeficiency syndrome; human immunodeficiency virus;
XX KW HIV-related disorder; differential expression; drug screening;
XX KW viral replication modulation; diagnosis; prognosis; predisposition;
XX KW anti-HIV; gene therapy; antisense therapy; human; TTK protein kinase;
XX KW enzyme.
XX OS Homo sapiens.
XX PN WO2003070883-A2.
XX PD 28-AUG-2003.
XX PF 13-FEB-2003; 2003WO-US004246.
XX PR 15-FEB-2002; 2002US-0357391P.
XX PR 13-MAY-2002; 2002US-0380249P.
XX PR 25-JUN-2002; 2002US-0391306P.
XX PR 27-AUG-2002; 2002US-0406297P.
XX PR 19-SEP-2002; 2002US-0412007P.
XX PR 10-OCT-2002; 2002US-0417508P.
XX PR 10-DEC-2002; 2002US-0432318P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Powell DM, Weich NS;
XX DR WPI; 2003-671808/63.
XX DR N-PSDB; ADE40436.

PT Identifying a compound capable of diagnosing, preventing or treating AIDS
PT or an HIV-related disorder comprises assaying the ability of the compound
PT to modulate e.g. 1414, 1481 or 1553 nucleic acid expression or
PT polypeptide activity.

PS Claim 1; SEQ ID NO 16; 167pp; English.

XX The invention relates to a method of identifying a compound useful in the
CC treatment of AIDS (acquired immunodeficiency syndrome) or an HIV (human
CC immunodeficiency virus)-related disorder. The invention involves assaying
CC the ability of a test compound to modulate the activity or expression of
CC 26 human proteins. These proteins and nucleic acids encoding them
CC (ADE40422-ADE40473) are differentially expressed in tissues relating to
CC AIDS or an HIV-related disorder compared to their expression in normal
CC tissues. The invention also relates to the use of the compounds
CC identified to modulate viral replication in a cell and to treat a patient
CC with AIDS or an HIV-related disorder. The invention further discloses
CC methods for the diagnostic evaluation and prognosis of various HIV-
CC related disorders, and for the identification of individuals exhibiting a
CC predisposition to such conditions. The modulatory compounds identified
CC using the method of the invention may be small organic molecules,
CC peptides, antibodies or antisense nucleic acid molecules. The methods of
CC the invention are useful in diagnosing, preventing or treating AIDS or
CC HIV-related disorders. The present sequence represents a human protein
CC which is differentially expressed in AIDS or HIV-related disorders.

XX SQ Sequence 841 AA;

Query Match 100.0%; Score 4376; DB 7; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.1e-292;

Matches		841;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MNKVRDIKNKFNEDLTDELSLNKISADTTDNSGTVNQIMMANNPEDWLSLLKLEKNS	60							
Db	1	MNKVRDIKNKFNEDLTDELSLNKISADTTDNSGTVNQIMMANNPEDWLSLLKLEKNS	60							
QY	61	VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR	120							
Db	61	VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR	120							
QY	121	ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNLNLQKKQLLS	180							
Db	121	ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNLNLQKKQLLS	180							
QY	181	EEKKNLSASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPDAEIGYR	240							
Db	181	EEKKNLSASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPDAEIGYR	240							
QY	241	NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP	300							
Db	241	NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP	300							
QY	301	SGNDSCELRNLKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESSLLAKLEETKEYQE	360							
Db	301	SGNDSCELRNLKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESSLLAKLEETKEYQE	360							
QY	361	PEVPESNQWQAKRKSECIQNQNPASSNHQIPELARKVNTQKHTTFFEQPVFSVSKQS	420							
Db	361	PEVPESNQWQAKRKSECIQNQNPASSNHQIPELARKVNTQKHTTFFEQPVFSVSKQS	420							
QY	421	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFFPACQLSTPYGQPACFQQQOH	480							
Db	421	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFFPACQLSTPYGQPACFQQQOH	480							
QY	481	QILATPLQNLQVLASSSANECISVKGRIYSILKQIGSGGSKVFQVLNEKKQIYAIKYVN	540							
Db	481	QILATPLQNLQVLASSSANECISVKGRIYSILKQIGSGGSKVFQVLNEKKQIYAIKYVN	540							
QY	541	LEEADNQTLDYSYRNEIAYLNKLQHSDKIIRLYDYEITDQIYVMMECGNIDLNLWLKXK	600							
Db	541	LEEADNQTLDYSYRNEIAYLNKLQHSDKIIRLYDYEITDQIYVMMECGNIDLNLWLKXK	600							
QY	601	KSIDPWKRSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDIT	660							
Db	601	KSIDPWKRSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDIT	660							
QY	661	SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSIGCILYMTYGTPTFPQ	720							
Db	661	SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSIGCILYMTYGTPTFPQ	720							
QY	721	IINQISKLHAIIDPNHEIEFPDIPKDLQDVLKCCLKRDPPKQRIPIPELLAHPYVQIQTH	780							
Db	721	IINQISKLHAIIDPNHEIEFPDIPKDLQDVLKCCLKRDPPKQRIPIPELLAHPYVQIQTH	780							
QY	781	PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLYEHYSGGSHNSSSKTTFEKKRGK	840							
Db	781	PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLYEHYSGGSHNSSSKTTFEKKRGK	840							
QY	841	K 841								
Db	841	X 841								

RESULT 5
ADE31645
ID ADE31645 standard; protein; 841 AA.
XX
AC ADE31645;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human 1682 protein #SEQ ID 2.
XX

KW Antiarteriosclerotic; cardiant; vasotropic; antiinflammatory;
KW thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;
KW cardiovascular; disorder; ischaemia; aortic bending;
KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;
KW angina; cardiomyopathy; cardiac death.
XX
OS Homo sapiens.
XX
PN WO2003065984-A2.
XX
PD 14-AUG-2003.
XX
PF 29-JAN-2003; 2003WO-US002571.
XX
PR 01-FEB-2002; 2002US-0353224P.
PR 15-MAR-2002; 2002US-0364529P.
PR 19-APR-2002; 2002US-0373861P.
PR 29-APR-2002; 2002US-0376287P.
PR 12-JUN-2002; 2002US-0388080P.
PR 24-JUN-2002; 2002US-0390971P.
PR 03-JUL-2002; 2002US-0394130P.
PR 10-JUL-2002; 2002US-0394797P.
PR 21-AUG-2002; 2002US-0404904P.
PR 23-AUG-2002; 2002US-0405450P.
PR 04-SEP-2002; 2002US-0408070P.
PR 06-NOV-2002; 2002US-0424300P.
PR 05-DEC-2002; 2002US-0431042P.
PR 05-DEC-2002; 2002US-0431079P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Logan TJ, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
PI Stagliano N, Perodin J, Rodrigue-Way A;
XX
DR WPI; 2003-731468/69.
DR N-PSDB; ADE31644.
XX
PT Identifying a compound capable of treating a cardiovascular disorder
PT (e.g. atherosclerosis) comprises assaying the ability of the compound to
PT modulate the expression or activity of e.g. 1682, 6169 or 6193
PT polypeptide or nucleic acid.
XX
PS Disclosure; SEQ ID NO 2; 328pp; English.
XX
CC The invention relates to a method for identifying a compound capable of
CC treating a cardiovascular disorder. The present invention identifies the
CC differential expression of 1682, 6169, 6193, 7771, 14395, 29002, 33216,
CC 43726, 69292, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,
CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,
CC 345, 9252, 9135, 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912,
CC 2868, 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,
CC 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or
CC 6585 genes in cardiovascular disease states. The methods are useful in
CC diagnosing, preventing and treating cardiovascular disorders, such as
CC atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury,
CC restenosis, arterial inflammation, vascular wall remodeling, aortic
CC microembolism, tachycardia, bradycardia, pressure overload, aortic
CC bending, coronary artery ligation, vascular heart disease, valvular
CC disease, including but not limited to, valvular degeneration caused by
CC calcification, rheumatic heart disease, endocarditis, or complications of
CC artificial valves; atrial fibrillation, long-QT syndrome, congestive
CC heart failure, sinus node dysfunction, angina, heart failure,
CC hypertension, atrial fibrillation, atrial flutter, pericardial disease,
CC including but not limited to, pericardial effusion and pericarditis;
CC cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic
CC cardiomyopathy, myocardial infarction, coronary artery disease, coronary
CC artery spasm, ischaemic disease, arrhythmia, sudden cardiac death and
CC cardiovascular developmental disorders. The methods may also be used for
CC identifying compounds that modulate cardiovascular disorders. Sequences
CC given in ADE31644-ADE31769 represent the genes and proteins that may be
CC regulated by a compound of the invention.
XX
SQ Sequence 841 AA;

Query Match		100.0%;	Score 4376;	DB 7;	Length 841;
Best Local Similarity		100.0%;	Pred. No. 2.1e-292;		
Matches 841;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MNKVRDIKKNFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS	60		
Db	1	MNKVRDIKKNFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS	60		
QY	61	VPLSDALLNKLIGRYSAIEALPDKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR	120		
Db	61	VPLSDALLNKLIGRYSAIEALPDKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR	120		
QY	121	ANCKKFAFVHISFAQFELSQGNVKKSKQLLOKAVERGAVPLEMLEIALRNLNKKQLLS	180		
Db	121	ANCKKFAFVHISFAQFELSQGNVKKSKQLLOKAVERGAVPLEMLEIALRNLNKKQLLS	180		
QY	181	EEKKNLSASTVLTAEQSFSGSLGHLQNRNNSCDRGQTTKARFLYGENMPPQDAEIGYR	240		
Db	181	EEKKNLSASTVLTAEQSFSGSLGHLQNRNNSCDRGQTTKARFLYGENMPPQDAEIGYR	240		
QY	241	NSLRQTNKTKQSCPPFGRVPVNLNNSPDCDKTDDSVVPCFMKQTSRSECRDLVVPQSKP	300		
Db	241	NSLRQTNKTKQSCPPFGRVPVNLNNSPDCDKTDDSVVPCFMKQTSRSECRDLVVPQSKP	300		
QY	301	SGNDSCELNKLSQVNSHFKEPLVSEKSSSELIITDSITLKNKTESLLAKLEETKEYQE	360		
Db	301	SGNDSCELNKLSQVNSHFKEPLVSEKSSSELIITDSITLKNKTESLLAKLEETKEYQE	360		
QY	361	PEVPESNOKQWQAKRKSEKINONPAASNHWQIPELARKVNTQKHTEQHPVFSVSKQS	420		
Db	361	PEVPESNOKQWQAKRKSEKINONPAASNHWQIPELARKVNTQKHTEQHPVFSVSKQS	420		
QY	421	PPISTKWFDPKSIKTPSSNTLDDYMSCFRTPVVKNDFFPACQLSTPYGQPACFQQQKH	480		
Db	421	PPISTKWFDPKSIKTPSSNTLDDYMSCFRTPVVKNDFFPACQLSTPYGQPACFQQQKH	480		
QY	481	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVN	540		
Db	481	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVN	540		
QY	541	LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIIRLYDYEITDQYIYVMMECGNIDLSWLKX	600		
Db	541	LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIIRLYDYEITDQYIYVMMECGNIDLSWLKX	600		
QY	601	KSIDPWERKSYWKMLEAVHTIHOHGIVHSDLKXANFLIVDGMKLIDFGIANQMPDIT	660		
Db	601	KSIDPWERKSYWKMLEAVHTIHOHGIVHSDLKXANFLIVDGMKLIDFGIANQMPDIT	660		
QY	661	SVVKDSQVGTVMYMPPEATKDMSSSRENGKSKISPKSDVWSLGCILYMTYKTPFQQ	720		
Db	661	SVVKDSQVGTVMYMPPEATKDMSSSRENGKSKISPKSDVWSLGCILYMTYKTPFQQ	720		
QY	721	IINQISKLHAIIDPNHEIEFPDPIPEKDLQDVLUKCLKRDPKQRIPIPELLAHPYVQIQTH	780		
Db	721	IINQISKLHAIIDPNHEIEFPDPIPEKDLQDVLUKCLKRDPKQRIPIPELLAHPYVQIQTH	780		
QY	781	PVNQAKGTTEEMKYVLQVLNPNLSILKAATLYEHYSGGESHSSSSKTFEKKRGK	840		
Db	781	PVNQAKGTTEEMKYVLQVLNPNLSILKAATLYEHYSGGESHSSSSKTFEKKRGK	840		
QY	841	K 841			
Db	841	K 841			

RESULT 6
ADE38369
ID ADE38369 standard; protein; 857 AA.
XX
AC ADE38369;
XX
DT 29-JAN-2004 (first entry)

XX	Human protein 1682 amino acid sequence.
DE	tumorigenic disorder; angiogenic disorder; aberrant gene expression;
XX	aberrant protein activity; cytostatic; antithyroid; antidiabetic;
KW	ophthalmological; cancer; breast cancer; colon cancer; lung cancer;
KW	prostatic cancer; Grave's disease; diabetic retinopathy; protein 1682.
XX	
OS	Homo sapiens.
XX	
PN	WO2003065006-A2.
XX	
PD	07-AUG-2003.
XX	
XX	30-JAN-2003; 2003WO-US002588.
PR	31-JAN-2002; 2002US-0353600P.
PR	15-MAR-2002; 2002US-0364517P.
PR	09-APR-2002; 2002US-0371075P.
PR	10-APR-2002; 2002US-0371507P.
PR	16-APR-2002; 2002US-0372984P.
PR	19-APR-2002; 2002US-0374194P.
PR	24-MAY-2002; 2002US-0382995P.
PR	31-MAY-2002; 2002US-0385023P.
PR	14-JUN-2002; 2002US-0388853P.
PR	17-JUN-2002; 2002US-0389395P.
PR	25-JUN-2002; 2002US-0391324P.
PR	15-JUL-2002; 2002US-0395944P.
PR	22-JUL-2002; 2002US-0397726P.
PR	13-AUG-2002; 2002US-0403046P.
PR	22-AUG-2002; 2002US-0405155P.
PR	27-AUG-2002; 2002US-0406361P.
PR	25-OCT-2002; 2002US-0421195P.
PR	12-NOV-2002; 2002US-0425456P.
PR	19-NOV-2002; 2002US-0427626P.
PR	10-DEC-2002; 2002US-0432122P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Hunter JJ, Macbeth KJ, Tsai F, Lesoon A, Lightcap ES;
PI	Williamson MW, Rudolph-Owen LA;
XX	
DR	WPI; 2003-646176/61.
DR	N-PSDB; ADE38368.
XX	
PT	Treating subject having tumorigenic disorder or angiogenic disorder
PT	caused by aberrant polypeptide e.g., N-formylpeptide receptor or nucleic
PT	acid, by administering a modulator.
XX	
PS	Disclosure; SEQ ID NO 30; 454pp; English.
XX	
CC	This invention relates to a novel method of treating a human subject
CC	having a tumorigenic disorder or angiogenic disorder, caused by aberrant
CC	gene expression or activity of an isolated protein, by administering a
CC	modulator. The modulator may have cytostatic, antithyroid, antidiabetic
CC	or ophthalmological activity. The method is useful for treating a subject
CC	having a tumorigenic or angiogenic disorder, in particular for treating
CC	cancer (for example breast cancer, colon cancer, lung cancer or prostatic
CC	cancer) and, for example, Grave's disease and diabetic retinopathy. The
CC	present sequence is the amino acid sequence of the novel isolated human
CC	protein 1682 of the invention.
XX	
SQ	Sequence 857 AA;

Query Match 100.0%; Score 4376; DB 7; Length 857;
Best Local Similarity 100.0%; Pred. No. 2.1e-292;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKKNFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 60
Db 17 MNKVRDIKKNFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 76

QY 61 VPLSDALLNKLIGRYSAIEALPDKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR 120

Db 77 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR 136
QY 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLOKKAVERGAVPLEMLEIALRNLNLOKKQLLS 180
Db 137 ANCKKFAFVHISFAQFELSQGNVKKSKQLLOKKAVERGAVPLEMLEIALRNLNLOKKQLLS 196
QY 181 EEEKKNLSASTVLTAEFSFSGSLGHLQNRNNSCDSRGQTTKAPFLYGENMPPQDAEIGYR 240
Db 197 EEEKKNLSASTVLTAEFSFSGSLGHLQNRNNSCDSRGQTTKAPFLYGENMPPQDAEIGYR 256
QY 241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 300
Db 257 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 316
QY 301 SGNDSCELRNLKSVQNSHFKEPLVSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 317 SGNDSCELRNLKSVQNSHFKEPLVSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE 376
QY 361 PEVPESNQWQAKRSEKINQNPAAASNHWQIPELARKVNTQKHTTFFEQPVFSVKQS 420
Db 377 PEVPESNQWQAKRSEKINQNPAAASNHWQIPELARKVNTQKHTTFFEQPVFSVKQS 436
QY 421 PPISTKWFDPKSICTPPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQPACFQQQH 480
Db 437 PPISTKWFDPKSICTPPSSNTLDDYMSCRTPVVKNDFFPACQLSTPYGQPACFQQQH 496
QY 481 QILATPLQNLQVLASSANECISVKGRISILKQIGSGSSKVFQVLNEKKQIYAIKYN 540
Db 497 QILATPLQNLQVLASSANECISVKGRISILKQIGSGSSKVFQVLNEKKQIYAIKYN 556
QY 541 LEEADNQTLDYSRNEIAYLNKLOQHSKIRLYDYEITDQIYVMWMECGNIDLSWLKKK 600
Db 557 LEEADNQTLDYSRNEIAYLNKLOQHSKIRLYDYEITDQIYVMWMECGNIDLSWLKKK 616
QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT 660
Db 617 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT 676
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQQ 720
Db 677 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQQ 736
QY 721 IINQISKLHAIIDPNEHIEFPDIPKDLQDVLKCLKRDPKQISIPPELLAHYPYQIOTH 780
Db 737 IINQISKLHAIIDPNEHIEFPDIPKDLQDVLKCLKRDPKQISIPPELLAHYPYQIOTH 796
QY 781 PVNQMAKGTTEEMKYVLGVLNLSNPSILKAAKTLYEHYSGSHNSSSSKTTFEKKRGK 840
Db 797 PVNQMAKGTTEEMKYVLGVLNLSNPSILKAAKTLYEHYSGSHNSSSSKTTFEKKRGK 856
QY 841 K 841
Db 857 K 857

RESULT 7
ABU56724

ID ABU56724 standard; protein; 857 AA.
XX
AC ABU56724;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polypeptide #317.
XX
KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
OS Unidentified.

XX WO200286443-A2.
PN
XX 31-OCT-2002.
PD
XX
PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Aziz N, Murray R;
PI
XX WPI; 2003-093161/08.
DR N-PSDB; ABX76453.
DR
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 27; Page 434; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC invention
XX
SQ Sequence 857 AA;
Query Match 99.9%; Score 4373; DB 6; Length 857;
Best Local Similarity 99.9%; Pred. No. 3.4e-292;
Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNKYRDIKNFKNEDLTDELNLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 60
Db 17 MNKYRDIKNFKNEDLTDELNLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 76
QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR 120
Db 77 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR 136
QY 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLOKKAVERGAVPLEMLEIALRNLNLOKKQLLS 180
Db 137 ANCKKFAFVHISFAQFELSQGNVKKSKQLLOKKAVERGAVPLEMLEIALRNLNLOKKQLLS 196
QY 181 EEEKKNLSASTVLTAEFSFSGSLGHLQNRNNSCDSRGQTTKAPFLYGENMPPQDAEIGYR 240
Db 197 EEEKKNLSASTVLTAEFSFSGSLGHLQNRNNSCDSRGQTTKAPFLYGENMPPQDAEIGYR 256
QY 241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 300
Db 257 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 316

QY	301	SGNDSCELRNLKSVQNSHFKEPLVDSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE	360
Db	317	SGNDSCELRNLKSVQNSHFKEPLVDSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE	376
QY	361	PEVPESNQWQAKRKSEKINQNPAASSNHWQIPELARKVNTQKHTTPEQPVFSVSKQS	420
Db	377	PEVPESNQWQAKRKSEKINQNPAASSNHWQIPELARKVNTQKHTTPEQPVFSVSKQS	436
QY	421	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPFPACQLSTPYGQPACFQQQKH	480
Db	437	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPFPACQLSTPYGQPACFQQQKH	496
QY	481	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVN	540
Db	497	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVN	556
QY	541	LEEADNQTLDSYRNEIAYLNKLQHSDDKIIRLYDYVEITDQYIYVMMECGNIDLSWLKXK	600
Db	557	LEEADNQTLDSYRNEIAYLNKLQHSDDKIIRLYDYVEITDQYIYVMMECGNIDLSWLKXK	616
QY	601	KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT	660
Db	617	KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT	676
QY	661	SVVKDSQVGTNYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYGTTPFQQ	720
Db	677	SVVKDSQVGTNYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYGTTPFQQ	736
QY	721	IINQISKLHAIIDPNHEIEFPDIPKDLQDVLCCLKRDPKQKORISIPPELLAHPYVQIQTH	780
Db	737	IINQISKLHAIIDPNHEIEFPDIPKDLQDVLCCLKRDPKQKORISIPPELLAHPYVQIQTH	796
QY	781	PVNQMAKGTTEMKYVLGQVLGNLSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK	840
Db	797	PVNQMAKGTTEMKYVLGQVLGNLSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK	856
QY	841	K 841	
Db	857	K 857	
RESULT 8			
ID	ADB80470		
XX	ADB80470	standard; protein; 857 AA.	
AC	ADB80470;		
DT	04-DEC-2003	(first entry)	
DE	Ovarian cancer-associated protein #17.		
KW	cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;		
KW	post-operative chemotherapy; radiation therapy; tumour prognosis;		
OS	pre-cancerous lesion detection.		
PN	Homo sapiens.		
PD	WO2002102235-A2.		
PF	27-DEC-2002.		
PR	18-JUN-2002; 2002WO-US019297.		
PR	18-JUN-2001; 2001US-0299234P.		
PR	27-AUG-2001; 2001US-0315287P.		
PR	05-SEP-2001; 2001US-0317544P.		
PR	13-NOV-2001; 2001US-0350666P.		
PR	12-APR-2002; 2002US-0372246P.		
PA	(EOSB-) EOS BIOTECHNOLOGY INC.		
PI	Mack DH, Gish KC;		

XX	WPI; 2003-167431/16.		
DR	N-PSDB; ADB80469.		
PT	Detecting an ovarian cancer-associated transcript in a cell from a		
PT	patient, comprises contacting a biological sample from the patient with a		
PT	polynucleotide that hybridizes to an ovarian cancer gene.		
XX	Claim 13; Page 287; 332pp; English.		
XX	The invention relates to a method of detecting an ovarian cancer-		
CC	associated transcript in a cell from a patient, by contacting a		
CC	biological sample from the patient with a polynucleotide that selectively		
CC	hybridizes to a sequence at least 80% identical to any of one of 80		
CC	nucleic acid sequences given in the specification. The method is useful		
CC	in diagnosing ovarian cancer and in identifying and using agents and/or		
CC	targets that inhibit ovarian cancer. The nucleic acid molecule,		
CC	polypeptide and the antibody may also be used in detecting ovarian		
CC	cancers, monitoring and early detection of relapse following treatment,		
CC	monitoring response to therapy, selecting patients for post-operative		
CC	chemotherapy or radiation therapy, in selecting mode of therapy,		
CC	determining tumour prognosis, early detection of pre-cancerous lesions,		
CC	and as vaccines. This sequence corresponds to one of the proteins used		
CC	for the detection method of the invention.		
XX	Sequence 857 AA;		
QY	Query Match	99.9%;	Score 4373; DB 7; Length 857;
Db	Best Local Similarity	99.9%;	Pred. No. 3.4e-292;
QY	Matches 840; Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
QY	1 MNKVRDIKNKFNEDLTDELSINKISADTTDNSGTVNQIMMANNPDMWLSLLKLEKNS	60	
Db	17 MNKVRDIKNKFNEDLTDELSINKISADTTDNSGTVNQIMMANNPDMWLSLLKLEKNS	76	
QY	61 VPLSDALLNKLIGRYSQAIELPPDKYQONESFARIQVFAELKAIQEPDDARDYFQMAR	120	
Db	77 VPLSDALLNKLIGRYSQAIELPPDKYQONESFARIQVFAELKAIQEPDDARDYFQMAR	136	
QY	121 ANCKKFAFVHISFAQFELSQGNVKKSKOLLQKAVERGAVPLEMLEIALNLNQLKQLLS	180	
Db	137 ANCKKFAFVHISFAQFELSQGNVKKSKOLLQKAVERGAVPLEMLEIALNLNQLKQLLS	196	
QY	181 EEEKKNLSASTVLTAEPSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR	240	
Db	197 EEEKKNLSASTVLTAEPSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR	256	
QY	241 NSLRQTNKTKQSCPPFGRVPVNLNNSPCDVKTDDSVVPCFMKQRTSRSECRDLVVPQSKP	300	
Db	257 NSLRQTNKTKQSCPPFGRVPVNLNNSPCDVKTDDSVVPCFMKQRTSRSECRDLVVPQSKP	316	
QY	301 SGNDSCELRNLKSVQNSHFKEPLVDSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE	360	
Db	317 SGNDSCELRNLKSVQNSHFKEPLVDSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE	376	
QY	361 PEVPESNQWQAKRKSEKINQNPAASSNHWQIPELARKVNTQKHTTPEQPVFSVSKQS	420	
Db	377 PEVPESNQWQAKRKSEKINQNPAASSNHWQIPELARKVNTQKHTTPEQPVFSVSKQS	436	
QY	421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPFPACQLSTPYGQPACFQQQKH	480	
Db	437 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPFPACQLSTPYGQPACFQQQKH	496	
QY	481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVN	540	
Db	497 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVN	556	
QY	541 LEEADNQTLDSYRNEIAYLNKLQHSDDKIIRLYDYVEITDQYIYVMMECGNIDLSWLKXK	600	
Db	557 LEEADNQTLDSYRNEIAYLNKLQHSDDKIIRLYDYVEITDQYIYVMMECGNIDLSWLKXK	616	
QY	601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT	660	

Db 617 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMCLKIDFGIANQMOPDIT 676

QY 661 SVVKDSQVGTVNYMPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYGTTPFQQ 720

Db 677 SVVKDSQVGTVNYMPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYGTTPFQQ 736

QY 721 IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCLXKRDPKQISIPPELLAHPYVQIQTH 780

Db 737 IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCLXKRDPKQISIPPELLAHPYVQIQTH 796

QY 781 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK 840

Db 797 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK 856

QY 841 K 841

Db 857 K 857

RESULT 9

ADC37171

ID ADC37171 standard; protein; 857 AA.

AC ADC37171;

XX 18-DEC-2003 (first entry)

DT Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 4.

XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;

KW cancer; infectious disease; bone disease; AIDS;

KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;

KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;

KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.

XX Homo sapiens.

OS WO2003048202-A2.

XX 12-JUN-2003.

PF 03-DEC-2002; 2002WO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

XX (ASAH) ASahi KASEI KK.

XX Matsuda A, Muramatsu S;

PI WPI; 2003-505282/47.

XX N-PSDB; ADC37170.

XX New purified protein that activates nuclear factor kappa B (NF-kappaB),

PT useful for treating inflammation, autoimmune diseases, cancers,

PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or

PT ischemic disorders.

XX Claim 1; SEQ ID NO 4; 938pp; English.

XX The present invention relates to novel proteins and their coding

CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-

CC kappaB). The proteins and their coding sequences are useful for treating

CC a disease associated with NF-kappaB activation, such as inflammation,

CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,

CC neurodegenerative diseases, or ischaemic disorders.

XX Sequence 857 AA;

Query Match 99.9%; Score 4373; DB 7; Length 857;

Best Local Similarity 99.9%; Pred. No. 3.4e-292;

Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 60

Db 17 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 76

QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQGNESFARIOVRFAELKAIQEPDDDARDYFQMAR 120

Db 77 VPLSDALLNKLIGRYSQAIEALPPDKYQGNESFARIOVRFAELKAIQEPDDDARDYFQMAR 136

QY 121 ANCKKFAFVHISPAQFELSOGNVKSKQLLOKAVERGAVPLEMLEIALRNLNKQKQLLS 180

Db 137 ANCKKFAFVHISPAQFELSOGNVKSKQLLOKAVERGAVPLEMLEIALRNLNKQKQLLS 196

QY 181 EEEKKNLSASTVLTAESEFSGSLGHLQNRNNSCDSRGQTTKARPLYGENMPPQDAEIGYR 240

Db 197 EEEKKNLSASTVLTAESEFSGSLGHLQNRNNSCDSRGQTTKARPLYGENMPPQDAEIGYR 256

QY 241 NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDDSVVPCFMKRQTSRSECRDLVPGSKP 300

Db 257 NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDDSVVPCFMKRQTSRSECRDLVPGSKP 316

QY 301 SGNDSCELNLSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESSLLAKLEETKEYQE 360

Db 317 SGNDSCELNLSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESSLLAKLEETKEYQE 376

QY 361 PEVPESNQKQWAKRSEKINQNPAASNHWQIPELARKVNTQKHTTTEQPVFSVKQS 420

Db 377 PEVPESNQKQWAKRSEKINQNPAASNHWQIPELARKVNTQKHTTTEQPVFSVKQS 436

QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPFPACOLSTPYGQPACFQQQ 480

Db 437 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPFPACOLSTPYGQPACFQQQ 496

QY 481 QILATPLQNLQVLASSANECISVKGRISILKQIGSGSSKVFQVLEKQIYAIKYVN 540

Db 497 QILATPLQNLQVLASSANECISVKGRISILKQIGSGSSKVFQVLEKQIYAIKYVN 556

QY 541 LEEADNQTLDSEYRNEIAYLNKLQOHSDKIIRLYDYEITDQIYVMMECGNIDLSWLKKK 600

Db 557 LEEADNQTLDSEYRNEIAYLNKLQOHSDKIIRLYDYEITDQIYVMMECGNIDLSWLKKK 616

QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMCLKIDFGIANQMOPDIT 660

Db 617 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMCLKIDFGIANQMOPDIT 676

QY 661 SVVKDSQVGTVNYMPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYGTTPFQQ 720

Db 677 SVVKDSQVGTVNYMPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYGTTPFQQ 736

QY 721 IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCLKRDPKQISIPPELLAHPYVQIQTH 780

Db 737 IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCLKRDPKQISIPPELLAHPYVQIQTH 796

QY 781 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK 840

Db 797 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLYEHYSGGESHNSSSSKTFEKKRGK 856

QY 841 K 841

Db 857 K 857

RESULT 10

ADC37169

ID ADC37169 standard; protein; 830 AA.

XX ADC37169;

AC ADC37169;

XX 18-DEC-2003 (first entry)

DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 2.

XX

KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; murine.
XX Mus musculus.
OS WO2003048202-A2.
XX 12-JUN-2003.
PF 03-DEC-2002; 2002WO-JP012644.
XX 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX (ASAH) ASahi KASEI KK.
PA Matsuda A, Muramatsu S;
XX WPI; 2003-505282/47.
XX N-PSDB; ADC37168.
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX Claim 1; SEQ ID NO 2; 938pp; English.
XX The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX Sequence 830 AA;
Query Match 72.7%; Score 3182.5; DB 7; Length 830;
Best Local Similarity 74.0%; Pred. No. 3.3e-210;
Matches 622; Conservative 87; Mismatches 105; Indels 27; Gaps 7;
QY 1 MNKVRDIKNKFNEDLTDELSLNKISADTTNSGTVNQIMMANPBDWLSLLKLEKNS 60
Db 17 MSKMRDIKNKI-NEDCTDELSLKICADHTE---TVNQIMRVGNTPENWLNFLKLEKNS 72
QY 61 VPLSDALLNKLIQRYSAIEALPPDKYQNESFARIQVRFAELKAIQEPDDARDYFQMAR 120
Db 73 SPLNDDLNLKLIQRYSAIEALPPDKYQNESFARIQVRFAELKAIQEPDDARDYFQMAR 132
QY 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLELALRNLNKQKQLLS 180
Db 133 ENCKKFAFVHVSFAQFELSQGNLKKSEQLLHKAVETGAVPLQMLETAMENHLQKQLLP 192
QY 181 EEEKNLSASTVLTAESEFSGSLGHLQNRNNSCDRGQTTKARFLYGENMPPQDAEIGYR 240
Db 193 EEDKKSASASTVLSAQEPFSSSLGNVQNRISCSERGAARVLYGENLPQDAEVRHQ 252
QY 241 NSLRQNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQRTSRSECRDLVVPGSKP 300
Db 253 NPFKQTHAAKRCPCFGRVPVNLNNSPFDYVKTDDSAVTLTRKGSQPD-RDAILPGSRP 311
QY 301 SGNDSCELRNLKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 312 RGSDSVELRGLKPIQTIYILKSLVSNKSSSEL-MSDLIALKSKTDS-LTKLEETK---- 365
QY 361 PEVPEGNQKQWQAKRKSECNQNPAASSNHWQIPELARKVNTQKHTTPEQPVFSVSKQS 420
Db 366 PEIAERPMQWQSTRKPECVFQNPAAFAPLRHVDVTPKAD-----KES 409

QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPFPACQLSTPYGQPACFQQQOH 480
Db 410 PPISVPKWLDPKSACETPSSSLDDYMKCFKTPVVKNDPFPACPSSTPYSQLARLQQQQQ 469
QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVN 540
Db 470 QGLSTPLQSLQISGSSSINECISVNGRIYSILKQIGSGSSKVFQVLNEKKQINAIKYVN 529
QY 541 LEEADNQTLDYRNEIAYLNKLQOHSDKIIRLYDYEITDQYIYVMVECGNIDLSWLKKK 600
Db 530 LEDADSTQIESYRNEIAFLNKLQOHSDKIIRLYDYEITEQYIYVMVECGNIDLSWLKKK 589
QY 601 KSIDPWERKSYWKMLAEVHTIHOHGIVHSDLKPANFLIVDGMKLKIDDFGIANQMQPDTT 660
Db 590 KSINPWERKSYWKMLAEVHTIHOHGIVHSDLKPANFVIDGMKLKIDDFGIANQMQPDTT 649
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYGTPTFQQ 720
Db 650 SIVKDSQVGTVMYMAPEAIRDMSSSRENSKIRTKVSPRSDVWSLGCILYYMTYGTPTFQH 709
QY 721 IINQISKLHAIIDPNHEIEFPDIPEKLDQVLKCCCLKRDPKQIRISIPELLAHYVQIQTH 780
Db 710 IINQVSKLHAIINPAHEIEFPDIPEKLDQVLKCCCLKRDPKQIRISIPELLTHPYVQIQPH 769
QY 781 PVNQMAKGTTEEMKYVLGQVLGNLSPNSILKAAKTLYEHYSGGSHNSSSSKTFEKKRGK 840
Db 770 PGSMARGATDEMKYVLGQVLGNLSPNSILKAAKTLYERINCGEQDSSSSKTFDKKRER 829
QY 841 K 841
Db 830 K 830
RESULT 11
ABP54948
ID ABP54948 standard; protein; 678 AA.
XX AC ABP54948;
XX DT 13-JAN-2003 (first entry)
XX Schizosaccharomyces pombe protein kinase Mph1.
DE Mph1; tyrosine threonine kinase; TTK; protein kinase; enzyme;
KW mitotic checkpoint; colon cancer; breast cancer; tumour; cytostatic;
KW gene therapy.
XX OS Schizosaccharomyces pombe.
XX WO200268444-A1.
PN 06-SEP-2002.
XX 21-FEB-2002; 2002WO-US005278.
PF 21-FEB-2001; 2001US-0271254P.
PR (CHIR) CHIRON CORP.
XX Reinhard C, Jefferson AB, Chan VW;
DR WPI; 2002-698650/75.
DR N-PSDB; ABV73994.
XX Reducing growth of cancer cells comprises reducing Tyrosine Threonine
PT Kinase (TTK) activity, useful in diagnosing and treating disorders with
PT abnormal expression levels and activity of TTK, such as lung, colon,
PT prostate and ovarian cancer.
XX Disclosure; Page 104-105; 113pp; English.
PS The present sequence is the protein sequence of Schizosaccharomyces pombe

QY	233	QDAEIGYRSLRQTNKTKQSCPFGRVPVNLNSPCDVKTDSDVVPFCMKRQTSRSECRD	292	PR	25-FEB-1999;	99US-0121825P.
Db	32	QDPELYFKN---DTFSKSS-----HSDGTVTGDTLRRQSSGATALE	70	PR	05-MAR-1999;	99US-0123180P.
QY	293	LVV--PGSK-----PSGNDSCELR--NLKSVQNSHFKEPLVSDKSSSELIITDSTLKNK	343	PR	09-MAR-1999;	99US-0123548P.
Db	71	RLVSHPTKNFDLQNGGGQNSALKVNTPAYQSMHFEHLITPLPS-----TNASH	121	PR	23-MAR-1999;	99US-0125788P.
QY	344	TESSLLAKLEETKEYQEPB-VPESNQK-----QWQAKRKSECINQNPAASSNHWOIPELA	397	PR	25-MAR-1999;	99US-0126264P.
Db	122	SEVLSAGVNDLNSNSEHDLPLKSVNKTPLGSLISRRRRIGRIGLGP-----PKRA	172	PR	25-FEB-2000;	2000EP-00301439.
QY	398	R-----KVNTQKHHTTPEQPVFSVSKQSPPISTS---KWFDPKSICKTPSSNITLDDYM	447	PR	25-FEB-1999;	99US-0121825P.
Db	173	EYTLTDPKSTDTKNSTEADEDIEMKSREVPASNVAATTLKPLQLHNTPLQTSQEHPK	232	PR	05-MAR-1999;	99US-0123180P.
QY	448	SCFRTPVVKNDPPACQL-----STPYGQP-ACFQQQQ-----HQI-----	482	PR	09-MAR-1999;	99US-0123548P.
Db	233	PSFHPSPQESSFSPRVQFDHVDERRASELHSPVTVFQEPQRSASQPYESHALSPKVAPL	292	PR	23-MAR-1999;	99US-0125788P.
QY	483	----LATPLQNLQVLASSANECISVKGRIVSILKQIGSGGSKVFQVLN-EKKQIYAIAK	537	PR	25-MAR-1999;	99US-0126264P.
Db	293	FDNSQATPIPKRQ-----QDVVTVANLOFIKLVGVGKGSSMVYRIFSPDNRSRLYALK	345	PR	25-FEB-2000;	2000EP-00301439.
QY	538	YVNLEEADNQTLDYSRNEIAYLNKLQHQHSDKIIRLYDYEITDQ--YIYMMECGNIDLNS	595	PR	25-FEB-1999;	99US-0121825P.
Db	346	EVNFINADQTTIQGYKNEIALRLK-SGNDRIIKLYAAEVNDTLGQLNMVMECGETDLAN	404	PR	05-MAR-1999;	99US-0123180P.
QY	596	WLKK--KKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLKIDFGIAN	653	PR	09-MAR-1999;	99US-0123548P.
Db	405	LLMKMKKPINLNFIRMYWEQMLEAVQVVDQNIHVHSDLKPANFLIVEGNLKLIDFGIAK	464	PR	23-MAR-1999;	99US-0125788P.
QY	654	QMOPDTTSVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYVMTY	713	PR	25-MAR-1999;	99US-0126264P.
Db	465	AIGNDTTNIHRDISHIGTINYMAPEALTDMNAHTNSGVKLKGRPSDVWSLGCILYQWVY	524	PR	25-FEB-2000;	2000EP-00301439.
QY	714	GKTPFQIINQISKLHAIIDPNHEIEFPD-----IPEK-----DLQDVLKCC	755	PR	25-FEB-1999;	99US-0121825P.
Db	525	GRAPFAH-LKMIQIAAIPNEQYHIFPEVALPANAVQEKESLPGVTGPDLMVVKRC	583	PR	05-MAR-1999;	99US-0123180P.
QY	756	LKRDPKQRISIPPELLAHPYVQIQTHPVNQMAK	787	PR	09-MAR-1999;	99US-0123548P.
Db	584	LERDQKRLLTIPELLVHPFLNPLPSYLTPLAK	615	PR	23-MAR-1999;	99US-0125788P.
RESULT 13						
AAG38500						
ID	AAG38500 standard; protein; 777 AA.					
XX	AAG38500;					
AC						
XX	18-OCT-2000 (first entry)					
DT						
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 47506.					
DE						
XX	Protein identification; signal transduction pathway; metabolic pathway;					
KW	hybridisation assay; genetic mapping; gene expression control; promoter;					
KW	termination sequence.					
XX	Arabidopsis thaliana.					
OS						
XX	EP1033405-A2.					
PN						
XX	06-SEP-2000.					
PD						
XX	25-FEB-2000; 2000EP-00301439.					
PF						
XX	25-FEB-1999; 99US-0121825P.					
XX	05-MAR-1999; 99US-0123180P.					
PR	09-MAR-1999; 99US-0123548P.					
PR	23-MAR-1999; 99US-0125788P.					
PR	25-MAR-1999; 99US-0126264P.					

PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	28-APR-1999;	99US-0130891P.
PR	30-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	04-MAY-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0135021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.

PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Query Match 14.7%; Score 643; DB 3; Length 720;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 28.1%; Pred. No. 2.3e-35;		
PR	05-AUG-1999;	99US-0147260P.	Matches 205; Conservative 99; Mismatches 242; Indels 184; Gaps 24;		
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160989P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			
PR	26-OCT-1999;	99US-0161359P.			
PR	26-OCT-1999;	99US-0161360P.			
PR	26-OCT-1999;	99US-0161361P.			
PR	28-OCT-1999;	99US-0161920P.			
PR	28-OCT-1999;	99US-0161992P.			
PR	28-OCT-1999;	99US-0161993P.			
QY	107	QEPDDARDYFQMARAN-----CKKFAFVHISFAQFELSQGNVYKSKQLQKAVERGAVP	160		
Db	30	QRSQDVVTLQSLSAANTLTQDTTNNLAITSVAGESASITQPTVSEHFNPSPDRQMDFGKSA	89		
QY	161	LEMLEIALRNLNLQKKQLLSEEEKKNLSASIVLTAE-----SFGSLGHLQNRNNSC-D	214		
Db	90	VTSLE---SNLDVQRK-----SQSLIGTSQDMEWDATNQAEASHL-----DACIG	131		
QY	215	SRGQTTKARFLYGENMPPQDAEIGYRNSLRQTNK-TKQSCPFGRVPVNNLLNSPDCDVKTD	273		
Db	132	SKHQ-----NLPVSDSEVSLKSEYKSSSLAKIQQQLGEPF-NFLNQP-----	173		
QY	274	DSVVPFCMKRQTSRSECRDLVVPGSKPSGNDSCELRNLSVQNSHFKEPLVSEKSSSELI	333		
Db	174	-----RTRC-----SAVGSSWATTTLI	190		
QY	334	ITDSITLKNKTESLLAKLEETKEYQEPVESPESNQKWQAKRKSECINQNPAASSNHWQI	393		
Db	191	HSSSAPMLNAT-----THVRSRYVEAD-----SNANPHAVQSQGNL	226		
QY	394	PELARKVNTQKHHTTPEQPVSFVSQSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTP	453		
Db	227	PSCC-----PSSKVSNIHLHPNKDATASEMPASTN--DPEVRVKETDTSKQQQITTGLEAP	279		
QY	454	VVKNDFPPACQLSTPYGPQ---ACFQQQHQILATPLQNL---QVLASSSANEC-----	501		
Db	280	VG-----SSIVGSDGQANARLPEELHTSVSSQPQKSDKHEKVASSKGPSAPRKR	329		
QY	502	-----ISVKGRIYSILKQIGSGSSKVFQVNLNEKKQIYAIKYVNLBEADNQTLDSYRNE	555		
Db	330	YDPLFFKVGKLYQRLGKIGSGSSEVHKVVISSDCTIYALKKIKLGRDYATAYGFCQE	389		
QY	556	IAYLNKLQHSDKIIRLYDYEITDQ-----YIYMMECGNIDL	593		
Db	390	IGYLLKLKGKTN-IIQLIDYEVTDKTLQLQEVLTGNTMSNKDGRVKEDGFIMVLEYGEIDL	448		
QY	594	NSWLKKK-----KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKL	645		
Db	449	AHMLSQKWRREIEGSDRTIDENWLRFYWQQILQAVNTIHEERIVHSDLKPANFLLVRGFLK	508		
QY	646	LIDFGIANQMOPDTTSVVKDSQVGTVNYMPPPEATKDMSSSRENGSKSKISPKSDVWSLG	705		
Db	509	LIDFGIAKAINSDTTNIQRDSQVGTLSYMSPEAFM-CNESDENGNT-IKGRPSDIWSLG	566		
QY	706	CILYMTYGTPTFQIINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLKRDPKQRI	765		
Db	567	CILYQMVYGRTPFADYKTFWAKFKVITDPNHEITYNQLSNPWLIDLMMKKCLAWDRNQWR	626		
QY	766	IPELLAHPYV	775		
Db	627	IPELLQHPFL	636		
RESULT 15					
AAG38502					
ID	AAG38502 standard; protein; 638 AA.				
XX					
AC	AAG38502;				
XX					
DT	18-OCT-2000 (first entry)				
XX					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 47508.				
XX					
KW	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence.				
XX					

OS	Arabidopsis thaliana.		
XX	EP1033405-A2.		
PN	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-00301439.		
PD	25-FEB-1999; 99US-0121825P.	30-JUN-1999;	99US-0141287P.
PF	05-MAR-1999; 99US-0123180P.	01-JUL-1999;	99US-0141842P.
XX	09-MAR-1999; 99US-0123548P.	01-JUL-1999;	99US-0142154P.
PR	23-MAR-1999; 99US-0125788P.	02-JUL-1999;	99US-0142055P.
PR	25-MAR-1999; 99US-0126264P.	06-JUL-1999;	99US-0142390P.
PR	29-MAR-1999; 99US-0126785P.	08-JUL-1999;	99US-0142803P.
PR	01-APR-1999; 99US-0127462P.	09-JUL-1999;	99US-0142920P.
PR	06-APR-1999; 99US-0128234P.	12-JUL-1999;	99US-0142977P.
PR	08-APR-1999; 99US-0128714P.	13-JUL-1999;	99US-0143542P.
PR	16-APR-1999; 99US-0129845P.	14-JUL-1999;	99US-0143624P.
PR	19-APR-1999; 99US-0130077P.	15-JUL-1999;	99US-0144005P.
PR	21-APR-1999; 99US-0130449P.	16-JUL-1999;	99US-0144085P.
PR	23-APR-1999; 99US-0130510P.	16-JUL-1999;	99US-0144086P.
PR	28-APR-1999; 99US-0131449P.	19-JUL-1999;	99US-0144325P.
PR	30-APR-1999; 99US-0132048P.	19-JUL-1999;	99US-0144331P.
PR	30-APR-1999; 99US-0132407P.	19-JUL-1999;	99US-0144332P.
PR	04-MAY-1999; 99US-0132484P.	19-JUL-1999;	99US-0144333P.
PR	05-MAY-1999; 99US-0132485P.	19-JUL-1999;	99US-0144334P.
PR	06-MAY-1999; 99US-0132486P.	19-JUL-1999;	99US-0144335P.
PR	06-MAY-1999; 99US-0132487P.	20-JUL-1999;	99US-0144352P.
PR	07-MAY-1999; 99US-0132863P.	20-JUL-1999;	99US-0144632P.
PR	11-MAY-1999; 99US-0134256P.	20-JUL-1999;	99US-0144632P.
PR	14-MAY-1999; 99US-0134218P.	20-JUL-1999;	99US-0144884P.
PR	14-MAY-1999; 99US-0134219P.	21-JUL-1999;	99US-0144814P.
PR	14-MAY-1999; 99US-0134221P.	21-JUL-1999;	99US-0145086P.
PR	14-MAY-1999; 99US-0134370P.	21-JUL-1999;	99US-0145088P.
PR	18-MAY-1999; 99US-0134768P.	22-JUL-1999;	99US-0145085P.
PR	19-MAY-1999; 99US-0134941P.	22-JUL-1999;	99US-0145087P.
PR	20-MAY-1999; 99US-0135124P.	22-JUL-1999;	99US-0145089P.
PR	21-MAY-1999; 99US-0135353P.	22-JUL-1999;	99US-0145192P.
PR	24-MAY-1999; 99US-0135629P.	23-JUL-1999;	99US-0145145P.
PR	25-MAY-1999; 99US-0136021P.	23-JUL-1999;	99US-0145218P.
PR	27-MAY-1999; 99US-0136392P.	23-JUL-1999;	99US-0145224P.
PR	28-MAY-1999; 99US-0136782P.	26-JUL-1999;	99US-0145276P.
PR	01-JUN-1999; 99US-0137222P.	27-JUL-1999;	99US-0145913P.
PR	03-JUN-1999; 99US-0137528P.	27-JUL-1999;	99US-0145918P.
PR	04-JUN-1999; 99US-0137502P.	27-JUL-1999;	99US-0145919P.
PR	07-JUN-1999; 99US-0137724P.	28-JUL-1999;	99US-0145951P.
PR	08-JUN-1999; 99US-0138094P.	02-AUG-1999;	99US-0146386P.
PR	10-JUN-1999; 99US-0138540P.	02-AUG-1999;	99US-0146388P.
PR	10-JUN-1999; 99US-0138847P.	02-AUG-1999;	99US-0146389P.
PR	14-JUN-1999; 99US-0139119P.	03-AUG-1999;	99US-0147038P.
PR	16-JUN-1999; 99US-0139452P.	04-AUG-1999;	99US-0147204P.
PR	16-JUN-1999; 99US-0139453P.	04-AUG-1999;	99US-0147302P.
PR	17-JUN-1999; 99US-0139492P.	05-AUG-1999;	99US-0147192P.
PR	18-JUN-1999; 99US-0139454P.	05-AUG-1999;	99US-0147260P.
PR	18-JUN-1999; 99US-0139455P.	06-AUG-1999;	99US-0147303P.
PR	18-JUN-1999; 99US-0139456P.	06-AUG-1999;	99US-0147416P.
PR	18-JUN-1999; 99US-0139457P.	09-AUG-1999;	99US-0147493P.
PR	18-JUN-1999; 99US-0139458P.	09-AUG-1999;	99US-0147935P.
PR	18-JUN-1999; 99US-0139459P.	10-AUG-1999;	99US-0148171P.
PR	18-JUN-1999; 99US-0139460P.	11-AUG-1999;	99US-0148319P.
PR	18-JUN-1999; 99US-0139462P.	12-AUG-1999;	99US-0148341P.
PR	18-JUN-1999; 99US-0139463P.	13-AUG-1999;	99US-0148565P.
PR	18-JUN-1999; 99US-0139461P.	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999; 99US-0139455P.	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999; 99US-0139456P.	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999; 99US-0139457P.	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999; 99US-0139458P.	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999; 99US-0139459P.	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999; 99US-0139460P.	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999; 99US-0139461P.	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999; 99US-0139462P.	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999; 99US-0139463P.	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999; 99US-0139763P.	26-AUG-1999;	99US-0150884P.
PR	21-JUN-1999; 99US-0139817P.	27-AUG-1999;	99US-0151065P.
PR	22-JUN-1999; 99US-0139899P.	27-AUG-1999;	99US-0151066P.
PR	23-JUN-1999; 99US-0140353P.	27-AUG-1999;	99US-0151080P.
PR	23-JUN-1999; 99US-0140354P.	30-AUG-1999;	99US-0151303P.
PR	24-JUN-1999; 99US-0140695P.	31-AUG-1999;	99US-0151438P.
PR	28-JUN-1999; 99US-0140823P.	01-SEP-1999;	99US-0151930P.
PR	29-JUN-1999; 99US-0140991P.	07-SEP-1999;	99US-0152363P.
		10-SEP-1999;	99US-0153070P.
		13-SEP-1999;	99US-0153758P.

```
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 14.6%; Score 640.5; DB 3; Length 638;
Best Local Similarity 29.2%; Pred. No. 2.8e-35;
Matches 187; Conservative 88; Mismatches 218; Indels 147; Gaps 18;

```
Qy 188 SASTVLTAESEFSGSLGHLQNRNNSCDSRGQTKARFLYG---ENMPPQDAEIGYRNSLR 244
Db 10 SLESLDVQRKSQSLIGTSQDMEDATNQAEBASHLDACIGSKHQLNLPVDSVSLKSEYK 69
Qy 245 QTNK-TKQSCPFGRVPVNLNNSPCDCVKTDDSVVPCFMKQTSRSECRDLVVPGSKPSGN 303
Db 70 DSSSLAKIQQLGEFP-NFLNQP-----RTRC----- 95
Qy 304 DSCELRNLSVQNSHFKPEPLVSDKSSSELIITDSITLKNKTESSILAKLEETKEYQPEV 363
Db 96 -----SAVGSSWATTTLIHSSAPMLNAT-----THVSRSYVEAD- 130
Qy 364 PESNQKQWQAKRKSECINQNPAASSNHWQIPELARKVNTQKHTTTEQPVFSVSKQSPPI 423
Db 131 -----SNANPHAVQSQGNLPSCC-----PSSKVSNIILHPNKDATASEMPA 170
Qy 424 STSKWFDPKSICTPSSNTLDDYMSCFRTPVVKNDPPACQLSTPYGQP-----ACFOQQQ 479
Db 171 STN---DPEVRVKETDTSKQQQITTGLEAPVG-----SSYGSDDGOANARLPEEL 217
Qy 480 HQILATPLQNL---QVLASSANEC-----ISVKGRIVSILKQIGSGGSKVFPQ 525
Db 218 HTSVSSQPQSKDKHEKVASSKGPSAPRKRNVDPDLPFKVNGKLYQRLGKIGSGGSSEVHK 277
Qy 526 VLNEKKQIYAIKYVNLBEADNQTLDSYRNEIAYLNKLQHQHSDKTIIRLYDYEITDQ----- 580
```

```
Db 278 VISSDCTIYALKKIKLGRDYATAYGFCQEIFGYLKKLGKTN-IIQLIDYEVTDKTLQE 336
Qy 581 -----YIYMMECGNIDINSWLKKK-----KSIDPWERKSYWKNM 615
Db 337 VLNGTMSNKDGRVKEDGFIYMVLEYGEIDLAAHMLSQKWREIEGSDRTIDENWLRFYWQOI 396
Qy 616 LEAVHTIHQHGIVHSDILKPANFLIVDGMKLKIDFGIANQMOPDTTSVVVKDSQVGTVNYMP 675
Db 397 LQAVNTIHEERIVHSDILKPANFLVVRGFLKIDFGIAKAINSDDTNIQRDSQVGTLSYMS 456
Qy 676 PEAKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYKTPFQQIINQISKLHAIIDPN 735
Db 457 PEAFM-CNESDENGNT-IKCGRPSDIWSLGCILYQMVYGRTPPFADYKTFWAKFKVITDPN 514
Qy 736 HEIEFPDIPEKDLOQVLKCCCLKRDPKQRIPIPELLAHPYV 775
Db 515 HEITYNQLSNPWLIDLMMKKCLAWDRNQRWRIPELLQHPFL 554
```

Search completed: September 29, 2004, 16:16:14
Job time : 141 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: September 29, 2004, 16:10:15 ; Search time 34 Seconds
(without alignments)
1276.984 Million cell updates/sec
Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDIKNKFKNEDLTDEL.....GGESHNSSSSKTFEKKRGKK 841
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	310.5	7.1	890	1 US-08-472-934-8	Sequence 8, Appli
2	310.5	7.1	890	2 US-08-323-460A-8	Sequence 8, Appli
3	310.5	7.1	890	2 US-08-461-146C-8	Sequence 8, Appli
4	310.5	7.1	890	3 US-08-461-145C-8	Sequence 8, Appli
5	310.5	7.1	890	4 US-08-628-829-12	Sequence 12, Appli
6	306	7.0	505	4 US-09-729-995-4	Sequence 4, Appli
7	306	7.0	505	4 US-10-135-689-4	Sequence 4, Appli
8	303.5	6.9	1247	1 US-08-472-934-10	Sequence 10, Appli
9	303.5	6.9	1247	2 US-08-323-460A-10	Sequence 10, Appli
10	303.5	6.9	1247	2 US-08-461-146C-10	Sequence 10, Appli
11	303.5	6.9	1247	3 US-08-461-145C-10	Sequence 10, Appli
12	303.5	6.9	1597	4 US-09-423-890-13	Sequence 13, Appli
13	303.5	6.9	1597	4 US-08-628-829-14	Sequence 14, Appli
14	298.5	6.8	505	4 US-09-729-995-2	Sequence 2, Appli
15	298.5	6.8	505	4 US-10-135-689-2	Sequence 2, Appli
16	294	6.7	279	4 US-09-799-875-23	Sequence 23, Appli
17	291	6.6	416	1 US-08-252-995D-2	Sequence 2, Appli
18	291	6.6	416	2 US-08-834-108-2	Sequence 2, Appli
19	291	6.6	464	1 US-08-252-995D-6	Sequence 6, Appli
20	291	6.6	464	2 US-08-834-108-6	Sequence 6, Appli
21	291	6.6	647	3 US-09-031-563-7	Sequence 7, Appli
22	291	6.6	647	4 US-09-392-277-7	Sequence 7, Appli
23	291	6.6	647	4 US-09-258-000-7	Sequence 7, Appli
24	291	6.6	648	3 US-09-031-563-5	Sequence 5, Appli
25	291	6.6	648	4 US-09-392-277-5	Sequence 5, Appli
26	291	6.6	648	4 US-09-258-000-5	Sequence 5, Appli
27	291	6.6	925	1 US-08-252-995D-4	Sequence 4, Appli

28	291	6.6	925	2 US-08-834-108-4	Sequence 4, Appli
29	291	6.6	1315	3 US-09-031-563-2	Sequence 2, Appli
30	291	6.6	1315	4 US-09-293-505-10	Sequence 10, Appli
31	291	6.6	1315	4 US-09-392-277-2	Sequence 2, Appli
32	291	6.6	1315	4 US-09-258-000-2	Sequence 2, Appli
33	290	6.6	260	3 US-09-031-563-24	Sequence 24, Appli
34	290	6.6	260	4 US-09-392-277-24	Sequence 24, Appli
35	290	6.6	260	4 US-09-258-000-24	Sequence 24, Appli
36	289.5	6.6	544	4 US-09-688-188B-95	Sequence 95, Appli
37	289.5	6.6	544	4 US-09-291-417D-95	Sequence 95, Appli
38	288	6.6	1315	3 US-09-031-563-25	Sequence 25, Appli
39	288	6.6	1315	4 US-09-392-277-25	Sequence 25, Appli
40	288	6.6	1315	4 US-09-258-000-25	Sequence 25, Appli
41	285	6.5	544	3 US-08-559-397A-29	Sequence 29, Appli
42	284	6.5	274	2 US-07-857-224B-29	Sequence 29, Appli
43	283.5	6.5	275	4 US-09-803-671B-5	Sequence 5, Appli
44	282.5	6.5	431	2 US-09-211-930-5	Sequence 5, Appli
45	282.5	6.5	431	3 US-09-340-993-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-472-934-8
; Sequence 8, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,405,941
; FILING DATE: 15-Apr-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,460
; FILING DATE: 14-Oct-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr. Esq.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: CPI-004DVCP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941

		INFORMATION FOR SEQ ID NO: 8:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 890 amino acids	
		TYPE: amino acid	
		TOPOLOGY: linear	
		MOLECULE TYPE: protein	
		FEATURE:	
		NAME/KEY: Xaa = Any amino acid	
		LOCATION: 116	
		US-08-472-934-8	
		Query Match	
		Best Local Similarity 7.1%; Score 310.5; DB 1; Length 890;	
		Matches 191; Conservative 150; Mismatches 307; Indels 299; Gaps 42;	
QY	39	IMMANNPEDWLSLLKLEKNSVPLSDAL----	LNKLIGRYSQAIEA----LPPDKYG- 88
Db	34	LMVYFDYMRSWIQLQLPQASHSLKNLLEEWNTKEITHYIRGGEAQAGKLCFCDIAGM 93	
QY	89	---QNESF--ARIQVRFAELKAIQEPDDARDYFQMARANCKKFAFVHISFAQFELSQGNV 143	
Db	94	LLKSTGSPLESGLQESCAELWTXADDNGAAD--ELRR-----SVIEIS----- 134	
QY	144	KSKQLQKAVERGAVPLEMLEIALRNLNQKQLLSEEEKKNLSASTVLTAESEFSGSL 203	
Db	135	RALKELFHEARERASKALGFAMLRKDLKLEIAAEFFVLSASARELLDA---LKAKQYVKVQI 191	
QY	204	GHQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYVNSLRQTNKTKQSCFFGRVPVNL 263	
Db	192	PGLEN-----LHVFVPDSLAEKK-----IILQLL 216	
QY	264	NS---PDCDVKTDDSVVPCFMKQTSRSECRDLVVPKSPGNDSCELRNLSKSV----- 314	
Db	217	NAATGKDCSKDPDDVFMDFALLTKHGDRARD-----SEDGWTWEARAVKIVPOVETV 270	
QY	315	-----QNSHKEPLVSDEKSSSELLITDSITLKNKTES--LLAK-LEET 355	
Db	271	DTLRSMQVNDNLLVVMESAH----LVLQRKAFQSIQGLMTVRHEQTSSQPIIAKGLQQL 326	
QY	356	K-----EYQEPVPEGNQKQWQAKRKSEGINQNPAASSNHWQ 392	
Db	327	KNDALELCNRISDAIDRVDMFTLEP-DAEVESESATLQYYREAMIQGYNFGFEYHKE 385	
QY	393	I-----PELARKVNTQKHTTPEQPVFSV-----SKQSPPISTSKWFD-----PKSI 434	
Db	386	VVRLMSGEFQKIG--DKYITPAQKWMNYVLTKECSGRGTRPRWATQGFDFLOAIEPAFI 443	
QY	435	CKTPSSNTLDDYMS-----CFRTPVKNDFPPAC---QLSTPYGQFACFQ--QQHQI 482	
Db	444	SALPE-----DDFLSLQALMNECIGHVIGKPHSPVTAIHRNSPRPVKPRCHSDPPNPHLI 499	
QY	483	LATP-----LQ-----NL 490	
Db	500	IPTPEGFRGSSVPENDRLASIAAELQFRSLRSHSSPTEERDEPAYPRSDSSGSTRRWEL 559	
QY	491	QVLASSANECISVKGRIVSILK----- 513	
Db	560	RTLISQF-KDSASKQGPPIEAQKSVRLFEERRYREMRKNIIGQVCDTPKSYDNVMHVGL 618	
QY	514	-----QIGSGGSSKVPQVNL-EKKQIVAIKVVNLEEADNQTLDSYRNEIAYLNK 561	
Db	619	RKVTFKQGNKIGEGQYGVYTCISVDITGELMAMKEIRFQPNDRHKTIKETADELKIFEG 678	
QY	562	LQHSKDIIRLYDYEITDQIYVWME-C--GNIDLSWLKKKSIDPWERKSYWKNMLEA 618	
Db	679	IKH--PNLVRYFGVELHREEMVIFMEYCDGTELEVSRLGLQEHV---IRLYTKQITVA 732	
QY	619	VHTIHOHGIVHSDLKPAN-FLIVDGMKLIDFGIANQMOPDITTSVVKD--SQVGTVMVMP 675	
Db	733	INVLHEGIVHRDIKGANIFLTSSGLIKLGDGFGSVKLKNAQTMPEGVNSTLGTAAAYMA 792	
QY	676	PEAIKDMSSSRENGKSKSPKSDVWSLGCILYVMTYGTPTPQQIINQISKLHAIDPN 735	
		Query Match	
		Best Local Similarity 7.1%; Score 310.5; DB 2; Length 890;	
		Matches 191; Conservative 150; Mismatches 307; Indels 299; Gaps 42;	
QY	39	IMMANNPEDWLSLLKLEKNSVPLSDAL----	LNKLIGRYSQAIEA----LPPDKYG- 88
Db	34	LMVYFDYMRSWIQLQLPQASHSLKNLLEEWNTKEITHYIRGGEAQAGKLCFCDIAGM 93	
QY	89	---QNESF--ARIQVRFAELKAIQEPDDARDYFQMARANCKKFAFVHISFAQFELSQGNV 143	
Db	94	LLKSTGSPLESGLQESCAELWTXADDNGAAD--ELRR-----SVIEIS----- 134	
QY	144	KSKQLQKAVERGAVPLEMLEIALRNLNQKQLLSEEEKKNLSASTVLTAESEFSGSL 203	
Db	135	RALKELFHEARERASKALGFAMLRKDLKLEIAAEFFVLSASARELLDA---LKAKQYVKVQI 191	
QY	204	GHQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYVNSLRQTNKTKQSCFFGRVPVNL 263	
Db	192	PGLEN-----LHVFVPDSLAEKK-----IILQLL 216	

QY	264	NS----	PDCDVKTDDSVVPCFMKQRKTSRSECRDLVVPGSKPSGNDSCELRNLSKV-----	314
Dd	217	NAATGDCSKDPDDVFMDAFLLLTKHGDARD-----	SEDGWGTWEARAVKIVPQVETV	270
QY	315	-----	QNSHFKEPLVSDEKSSELIIITDSITLKNKTESS--LIAK-LEET	355
Dd	271	DTLRSMQVDNLLLVMESAH---	LVLQRKAQQSIEGLMTVRHEQTSSQPILAKGLQQL	326
QY	356	K-----	EYQEPEVPESNQKWQAARKSECINQNPAASNHHWQ	392
Dd	327	KNDALELCNRISDAIDRVDMFTLEF-DAEVEESESATLQQYYREAMIQYNFGFEYHKE	385	
QY	393	I-----	PELARKVNTEQKHTTFEQPVFSV-----SKOSPPISTSKWFD-----PKSI	434
Dd	386	VVRLMSGEFRQKIG--DKYIITFAQKMWNVYLTKCESGRTRPRWATQGDFLOAIEPAPI	443	
QY	435	CKTPSSNTLDDYMS-----	CFRTPVVKNDFFPAC---QLSTPYGQPACFQQ--QQHQI	482
Dd	444	SALPE---	DDFLSLQALMNECIGHVIGKPHSPVTAIHRNSRPVKVPRCHSDPENPHLI	499
QY	483	LATP-----	LQ-----NL	490
Dd	500	IPTPEGFRGSSVPENDRLASIAABELQFRSLSRHSSPTEERDEPAYPRSDSGSTRSWEL	559	
QY	491	QVLASSANECISVKGRIYSILK-----	-----	513
Dd	560	RTLISQT-KDSASKQGPIEAIOKSVRLFEERRYREMRKKNIIGQVCDTTPKSYDNYMHVGL	618	
QY	514	-----	QIGSGGSSKFQVLN-EKKOYIAKYVNLHEADNQTLDSYRNETAYLNK	561
Dd	619	RKVTFKWRGNKIGEGQYGVKYTCISVDTGELMAMKEIRFPNDHKTIKETADELKIFEG	678	
QY	562	LQHSDKIIRLYDYEITDQYIYMVE-C--GNIDLNSWLKKKSIDPWKRKSYWKMLSEA	618	
Dd	679	IKH--PNLVRYFGVELHREEMYIFNEYCDEGTLEEVSRLGLEHV----IRLYTKQITVA	732	
QY	619	VHTIHQHGLVHSDLKPAN-FLIVDGMLKLIDFGIANQMOPDTTSVVKD--SOVGTVNYMP	675	
Dd	733	INVLHEHGIVHRDIKGANIFLTSSGLIKLGDFGCSSVKLKNNAQTMPGEVNSTLGTAAAYMA	792	
QY	676	PEAKDMSSSRENGKSKISKIPKSDVWSLGCILYMYTYGTFPFQOIINOISKULHAIDPN	735	
Dd	793	PEVI-----	TRAKGEGHGR---AADISWLGCVVIEMVTGKRPWHYEHNFOIMYK-VGMG	843
QY	736	HEIEFPDPEK----	DLQDVLKCCLXRDPKQORISIPELLAHPYVQIQT	779
Dd	844	HK---PPIPELSPGKAFLSHCLESPPKIRWTASQLLDHAFVKVCT	887	

RESULT 3

US-08-461-146C-8
; Sequence 8, Application US/08461146C
; Patent No. 5981265
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,146C
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435

Db 145 AYNEREDRHYAMKVLKSKLLKQYGFRRPPRRGSAQGGPAKQLLPLERVYQEIAILK 204
QY 561 KLQK-HSDKIIRLYDYEITDQYIYMWMECGNIDLSWLKK-----KKSIDPWERKSY 611
Db 205 KLDHVNVLKIEVLDDPAEDN-LYLVPDL-----LRKGPVMEVPCDKPFPPEEQARLY 255
QY 612 WKNMLEAVHTIHQHGIVHSDLKIPANFLI-VDGMLKLIDFGIANQMOPDTTSVWKDSQVGT 670
Db 256 LRDIILGLEVLHCQKIVHRDIKPSNLLGDDGHVKIADFVSNQFEGNDAQL--SSTAGT 313
QY 671 VNYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQIINQISKLHA 730
Db 314 PAFMAPEAISTGQS-FSGKA-----LDVWATGVTLYCFVYGKCPF--IDEXILALHR 363
QY 731 IIDPNHEIEFPDPIE--KDLQVLKCKLKRDPKQKORISIPPELLAHPYV 775
Db 364 KI-KNEAVVPPEEPEVSEELKDLILKMLDKNPETRIGVSDIKLHPWV 409

RESULT 7
US-10-135-689-4
; Sequence 4, Application US/10135689
; Patent No. 6670162
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904DIV
; CURRENT APPLICATION NUMBER: US/10/135,689
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/247,031
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: 09/729,995
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-135-689-4

Query Match 7.0%; Score 306; DB 4; Length 505;
Best Local Similarity 25.1%; Pred. No. 2.9e-16;
Matches 117; Conservative 73; Mismatches 155; Indels 122; Gaps 21;
QY 349 LAKLEETKEYQEPEVPESNQKQWAKRKSECIQNQPAASSNHWQIPELARKVUNTEQKHTT 408
Db 25 VAHLEEAEEGPEP-----ASNGVDPPIPRARAASVIPGSAS 59
QY 409 FEQPVFSVSKSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPACQLSTP 468
Db 60 RPTPV-----RPSLSARKF---SLQERPAG-----SCLEAQVGPYSTGPASHMS-- 100
QY 469 YGQPACFQ---QQHQILATPLQNLQVLASSSANEISVKGRYISILKIGSGGSKVFQ 525
Db 101 ---PRAWRRPTIESHV-----AISTDECVQLNQ--YKLSQSEIGKGAYGVVRL 144
QY 526 VLNEKK-QIYAIKYVNLEADNQ-----TLDSEYRNEIAYLN 560
Db 145 AYNEREDRHYAMKVLKSKLLKQYGFRRPPRRGSAQGGPAKQLLPLERVYQEIAILK 204
QY 561 KLQK-HSDKIIRLYDYEITDQYIYMWMECGNIDLSWLKK-----KKSIDPWERKSY 611
Db 205 KLDHVNVLKIEVLDDPAEDN-LYLVPDL-----LRKGPVMEVPCDKPFPPEEQARLY 255
QY 612 WKNMLEAVHTIHQHGIVHSDLKIPANFLI-VDGMLKLIDFGIANQMOPDTTSVWKDSQVGT 670
Db 256 LRDIILGLEVLHCQKIVHRDIKPSNLLGDDGHVKIADFVSNQFEGNDAQL--SSTAGT 313
QY 671 VNYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQIINQISKLHA 730

Db 314 PAFMAPEAISTGQS-FSGKA-----LDVWATGVTLYCFVYGKCPF--IDEXILALHR 363
QY 731 IIDPNHEIEFPDPIE--KDLQVLKCKLKRDPKQKORISIPPELLAHPYV 775
Db 364 KI-KNEAVVPPEEPEVSEELKDLILKMLDKNPETRIGVSDIKLHPWV 409

RESULT 8
US-08-472-934-10
; Sequence 10, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,405,941
; FILING DATE: 15-Apr-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,460
; FILING DATE: 14-Oct-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr. Esq.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: CPI-004DVCP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-934-10

Query Match 6.9%; Score 303.5; DB 1; Length 1247;
Best Local Similarity 22.2%; Pred. No. 1.7e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
QY 296 PGSKPSGNDSCELRNLSKVONSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLE-- 353
Db 844 PG--PGGGDSVPAKPVNT-----APDTRGSSVPENDRL-----ASIAAELQFR 884

QY 354 -----ETKEYQEPVPESN-----QKQWQAKRKSECINQNPAASSNHWQIPELARKVN 401
Db 885 SLSRHSSPTEERDEPAYPRSDSSGSTRRSWELR---TLISQTKDSASKQGPIEAIOKSVR 941
QY 402 TEQKHTTFEQPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPP 461
Db 942 -----LFEERRYREMRKNIIG-----QVCDTPKS--YDNVMHVGLRKVT----- 979
QY 462 ACQLSTPYGQPACFOQOQHILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSS 521
Db 980 -----FKWQRG-----NKIGEGQYG 994
QY 522 KVFQVLN-EKKQIYAIKYVNLEEADNQTLDYSRNEIAYLNKLQHSDKIIRLYDYEITDQ 580
Db 995 KVTICISVDTGELMAMKEIRFPQNDHKTIKETADELKIFEGIKH--PNLVRYFGVELHRE 1052
QY 581 YIYMVME-C--GNIDLNSWLKKSIDPWERKSYWKNMLEAVHTIHOHGIVHSDLKPAN- 636
Db 1053 EMYIFMEYCDGTELEEVSRGLQEHV----IRLYTKQITVAINVLHEHGIVHRDIKGANI 1108
QY 637 FLIVDGMKLIDFGIANQMPDPTTSVVKD--SQVGTVNYMPPBRAIKDMSSSRENGSKSK 694
Db 1109 FLTSSGLIKLGFSGSVKLKNNAQTMPGEVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1163
QY 695 ISPKSDVWSLGCILYMTYGTPTFQIINQISKLHAIIDPNHEIEFPDPIEK---DLQDV 751
Db 1164 ---AADIWSLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKAF 1216
QY 752 LKCLKRDPKQKORISIPPELLAHPYVQIQT 779
Db 1217 LSHCLESDPKIRWTASQLLDHAFVKVCT 1244

RESULT 9

US-08-323-460A-10
; Sequence 10, Application US/08323460A
; Patent No. 5854043
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHERIDAN ROSS & MCINTOSH
; STREET: 1700 LINCOLN STREET, SUITE 3500
; CITY: DENVER
; STATE: CO
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,460A
; FILING DATE: 14-OCT-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,254
; FILING DATE: 14-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOVARIK, JOSEPH E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1247 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-323-460A-10
Query Match 6.9%; Score 303.5; DB 2; Length 1247;
Best Local Similarity 22.2%; Pred. No. 1.7e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
QY 296 PGSKPSGNDSCELRNLSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLE-- 353
Db 844 PG--PGGDSVPKPVNT-----APTRGSSVPENDRL-----ASIAAELOQR 884
QY 354 -----ETKEYQEPVPESN-----QKQWQAKRKSECINQNPAASSNHWQIPELARKVN 401
Db 885 SLSRHSSPTEERDEPAYPRSDSSGSTRRSWELR---TLISQTKDSASKQGPIEAIOKSVR 941
QY 402 TEQKHTTFEQPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPP 461
Db 942 -----LFEERRYREMRKNIIG-----QVCDTPKS--YDNVMHVGLRKVT----- 979
QY 462 ACQLSTPYGQPACFOQOQHILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSS 521
Db 980 -----FKWQRG-----NKIGEGQYG 994
QY 522 KVFQVLN-EKKQIYAIKYVNLEEADNQTLDYSRNEIAYLNKLQHSDKIIRLYDYEITDQ 580
Db 995 KVTICISVDTGELMAMKEIRFPQNDHKTIKETADELKIFEGIKH--PNLVRYFGVELHRE 1052
QY 581 YIYMVME-C--GNIDLNSWLKKSIDPWERKSYWKNMLEAVHTIHOHGIVHSDLKPAN- 636
Db 1053 EMYIFMEYCDGTELEEVSRGLQEHV----IRLYTKQITVAINVLHEHGIVHRDIKGANI 1108
QY 637 FLIVDGMKLIDFGIANQMPDPTTSVVKD--SQVGTVNYMPPBRAIKDMSSSRENGSKSK 694
Db 1109 FLTSSGLIKLGFSGSVKLKNNAQTMPGEVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1163
QY 695 ISPKSDVWSLGCILYMTYGTPTFQIINQISKLHAIIDPNHEIEFPDPIEK---DLQDV 751
Db 1164 ---AADIWSLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKAF 1216
QY 752 LKCLKRDPKQKORISIPPELLAHPYVQIQT 779
Db 1217 LSHCLESDPKIRWTASQLLDHAFVKVCT 1244

RESULT 10

US-08-461-146C-10
; Sequence 10, Application US/08461146C
; Patent No. 5981265
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,146C
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/354,516
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,254
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: KARA, Catherine J.
REGISTRATION NUMBER: P41,106
REFERENCE/DOCKET NUMBER: CPI-004CN3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-146C-10

Query Match 6.9%; Score 303.5; DB 2; Length 1247;
Best Local Similarity 22.2%; Pred. No. 1.7e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;

QY 296 PGSKPSGNDSCELRNKSVQNSHFKEPLVDSDEKSSSELIITDSITLKNKTESSLLAKLE-- 353
Db 844 PG--PGGGDSVPKPVNT-----APDTRGSSVPENDRL-----ASIAAELQFR 884

QY 354 -----ETKEYQEPEVPESN-----QKQWQAKRKSECINQNPAASSNHQIPELARKVN 401
Db 885 SLRSHSSPTERDEPAYPRSDSGSTRRSWELR---TLISQTKDSASKQGPPIEAIQKSVR 941

QY 402 TEQKHTTFEPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPP 461
Db 942 -----LFEERRYREMRKNIIG-----QVCDTPKS--YDNVMHVGLRKVT----- 979

QY 462 ACQLSTPYGQFACFQQQHQILATPLQNLQVLASSANECISVKGRISILKQIGSGSS 521
Db 980 -----FKWQRG-----NKIGEGQYG 994

QY 522 KVFQVLN-EKKQIYAIKYVNLEADNQTLDYSRNEIAYLNKLOQHSKIRLYDYEITDQ 580
Db 995 KVTCTISVDTGELMAMKEIRFQNDHKTIKETADELKIFEGIKH--FNLVRYFGVELHRE 1052

QY 581 YIYVME-C--GNIDLNSWLKSKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPAN- 636
Db 1053 EMYIFMEYDEGTLEEVSRGLQEHV---IRLYTKQITVAINVLEHGVHVDIKGANI 1108

QY 637 FLIVDGLMLKIDFGIANQMQPDTTSVWKD--SQVGTVMYMPPEAIKDMSSSRENGKSKS 694
Db 1109 FLTSSGLIKLDFGCSVKLNNAQTMPGEVNSTLGAAYMAPEVI-----TRAKGEGHGR 1163

QY 695 ISPKSDVWSLGCILYMTYGTPTFQQIINQISKLHAIDPNHEIEFPDIPEK--DLQDV 751
Db 1164 --AADIWSLGCVVIEMVTGKRPWHEYEHNFQIMYK-VGMGHK--PPIPERLSPEGKAF 1216

QY 752 LKCLKRDPKORISIPPELLAHPYVQIQT 779
Db 1217 LSHCLESDPKIRWTASQLLDHAFVKVCT 1244

RESULT 11
US-08-461-145C-10
; Sequence 10, Application US/08461145C
; Patent No. 6074861
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: NOVEL MEKK PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive and Cockfield, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,145C
FILING DATE: 5-JUNE-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,254
FILING DATE: 11-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/354,516
FILING DATE: 21-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: P-41,106
REFERENCE/DOCKET NUMBER: CPI-004CN1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-145C-10

Query Match 6.9%; Score 303.5; DB 3; Length 1247;
Best Local Similarity 22.2%; Pred. No. 1.7e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;

QY 296 PGSKPSGNDSCELRNKSVQNSHFKEPLVDSDEKSSSELIITDSITLKNKTESSLLAKLE-- 353
Db 844 PG--PGGGDSVPKPVNT-----APDTRGSSVPENDRL-----ASIAAELQFR 884

QY 354 -----ETKEYQEPEVPESN-----QKQWQAKRKSECINQNPAASSNHQIPELARKVN 401
Db 885 SLRSHSSPTERDEPAYPRSDSGSTRRSWELR---TLISQTKDSASKQGPPIEAIQKSVR 941

QY 402 TEQKHTTFEPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPP 461
Db 942 -----LFEERRYREMRKNIIG-----QVCDTPKS--YDNVMHVGLRKVT----- 979

QY 462 ACQLSTPYGQFACFQQQHQILATPLQNLQVLASSANECISVKGRISILKQIGSGSS 521
Db 980 -----FKWQRG-----NKIGEGQYG 994

QY 522 KVFQVLN-EKKQIYAIKYVNLEADNQTLDYSRNEIAYLNKLOQHSKIRLYDYEITDQ 580
Db 995 KVTCTISVDTGELMAMKEIRFQNDHKTIKETADELKIFEGIKH--FNLVRYFGVELHRE 1052

QY 581 YIYVME-C--GNIDLNSWLKSKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPAN- 636

Db 1053 EMYIFMEYCDGTLBEEVSRGLQEHV-----IRLYTKQITVAINVLHEHGIVHRDIKGANI 1108
Qy 637 FLIVDGMKLIDFGIANQMOPDTTSSVVKD--SQVGTVNYMPPPEAIKDMSSSRENGKSKS 694
Db 1109 FLTSSGLIKLGDGCSVKLKNNNAQTMPEGVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1163
Qy 695 ISPKSDVWSLGCILYYMTYGTPTFQIINQISKLHAIIDPNHEIEFFDPIEK---DLQDV 751
Db 1164 ---AADIWSLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKAF 1216
Qy 752 LKCCLRDPKQKORISIPPELLAHPPYVQIQ 779
Db 1217 LSHCLESDPKIRWTASQLLDHAFVKVCT 1244

RESULT 12
US-09-423-890-13
; Sequence 13, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEKK PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: CPI-085CPPC
; CURRENT APPLICATION NUMBER: US/09/423,890
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-423-890-13

Query Match 6.9%; Score 303.5; DB 4; Length 1597;
Best Local Similarity 22.2%; Pred. No. 2.5e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
Qy 296 PGSKPSGNDSCELRNLSVQNSHFKEPLVSDKSSSELIITDSTLKNKTESSLLAKLE-- 353
Db 1194 PG--PGGGDSVPAKPVNT-----APDTRGSSVPENDRL-----ASIAAELQFR 1234
Qy 354 -----ETKEYQEPVPEPN-----QKQWQAKRKSECIINQNPAASSNHQIPELARKVN 401
Db 1235 SLSRHSSPTEERDEPAYPRSDSSGSTRRWELR--TLISQTKDSASKQGPPIAIOKSVR 1291
Qy 402 TEQKHTTFEQPVFSVSKSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFFPP 461
Db 1292 -----LFEERRYREMRKNIIG-----QVCDTPKS--YDNVMHVGLRKVT----- 1329
Qy 462 ACQLSTPYGQPACFQQQHQILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSS 521
Db 1330 -----FKWQRG-----NKIGEGQYG 1344
Qy 522 KVFQVLN-EKKQIYAIKYVNLEEDNQTLDSYRNEIAYLNKLOQHSCKIIRLYDYEITDQ 580
Db 1345 KYVTCISVDTGELMAMKEIRFPNDHKTIKETADELKIFEGIKH--PNLVRYFGVELHRE 1402
Qy 581 YIYMVME-C--GNIDLNSWLKKKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPAN- 636
Db 1403 EMYIFMEYCDGTLBEEVSRGLQEHV-----IRLYTKQITVAINVLHEHGIVHRDIKGANI 1458
Qy 637 FLIVDGMKLIDFGIANQMOPDTTSSVVKD--SQVGTVNYMPPPEAIKDMSSSRENGKSKS 694
Db 1459 FLTSSGLIKLGDGCSVKLKNNNAQTMPEGVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1513
Qy 695 ISPKSDVWSLGCILYYMTYGTPTFQIINQISKLHAIIDPNHEIEFFDPIEK---DLQDV 751

Db 1514 ---AADIWSLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKAF 1566
Qy 752 LKCCLRDPKQKORISIPPELLAHPPYVQIQ 779
Db 1567 LSHCLESDPKIRWTASQLLDHAFVKVCT 1594
RESULT 13
US-08-628-829-14
; Sequence 14, Application US/08628829A
; Patent No. 6333170
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External
; FILE REFERENCE: CPI-004DVCP3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; CURRENT FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
; EARLIER APPLICATION NUMBER: 08/472,934
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-628-829-14

Query Match 6.9%; Score 303.5; DB 4; Length 1597;
Best Local Similarity 22.2%; Pred. No. 2.5e-15;
Matches 113; Conservative 99; Mismatches 165; Indels 131; Gaps 23;
Qy 296 PGSKPSGNDSCELRNLSVQNSHFKEPLVSDKSSSELIITDSTLKNKTESSLLAKLE-- 353
Db 1194 PG--PGGGDSVPAKPVNT-----APDTRGSSVPENDRL-----ASIAAELQFR 1234
Qy 354 -----ETKEYQEPVPEPN-----QKQWQAKRKSECIINQNPAASSNHQIPELARKVN 401
Db 1235 SLSRHSSPTEERDEPAYPRSDSSGSTRRWELR--TLISQTKDSASKQGPPIAIOKSVR 1291
Qy 402 TEQKHTTFEQPVFSVSKSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFFPP 461
Db 1292 -----LFEERRYREMRKNIIG-----QVCDTPKS--YDNVMHVGLRKVT----- 1329
Qy 462 ACQLSTPYGQPACFQQQHQILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSS 521
Db 1330 -----FKWQRG-----NKIGEGQYG 1344
Qy 522 KVFQVLN-EKKQIYAIKYVNLEEDNQTLDSYRNEIAYLNKLOQHSCKIIRLYDYEITDQ 580
Db 1345 KYVTCISVDTGELMAMKEIRFPNDHKTIKETADELKIFEGIKH--PNLVRYFGVELHRE 1402
Qy 581 YIYMVME-C--GNIDLNSWLKKKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPAN- 636
Db 1403 EMYIFMEYCDGTLBEEVSRGLQEHV-----IRLYTKQITVAINVLHEHGIVHRDIKGANI 1458
Qy 637 FLIVDGMKLIDFGIANQMOPDTTSSVVKD--SQVGTVNYMPPPEAIKDMSSSRENGKSKS 694
Db 1459 FLTSSGLIKLGDGCSVKLKNNNAQTMPEGVNSTLGTAAAYMAPEVI-----TRAKGEHGR 1513
Qy 695 ISPKSDVWSLGCILYYMTYGTPTFQIINQISKLHAIIDPNHEIEFFDPIEK---DLQDV 751
Db 1514 ---AADIWSLGCVVIVMTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKAF 1566
Qy 752 LKCCLRDPKQKORISIPPELLAHPPYVQIQ 779

Db 1567 LSHCLESDBKIRWTASQLLDHAPVKVCT 1594

RESULT 14

US-09-729-995-2

; Sequence 2, Application US/09729995

; Patent No. 6426206

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL000904

; CURRENT APPLICATION NUMBER: US/09/729,995

; CURRENT FILING DATE: 2000-12-06

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 505

; TYPE: PRT

; ORGANISM: Human

US-09-729-995-2

Query Match 6.8%; Score 298.5; DB 4; Length 505;

Best Local Similarity 28.8%; Pred. No. 1.2e-15;

Matches 92; Conservative 58; Mismatches 109; Indels 61; Gaps 14;

QY 493 LASSANECISVKGRISILKQIGSGSKVQVFLNEKK-QIYAIKYNLEADNQ----- 547

Db 114 VAISDAEDCVQLNQ--YKLQSEIGKGGAYGVVRLAYNESEDRHYAMKVLKSKKLLKQYGF 171

QY 548 -----TLDSEYRNEIAYLNKLQ--HSDKIIRLYDYEITDQYIYVM 586

Db 172 RPPPPRGSAAGGPAKQLPLERVYQEIALLKLDHVNVLKIEVLDDPAEDN-LYLVF 230

QY 587 ECGNIDLNSWLKK-----KKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFL 638

Db 231 DL-----LRKGPVMEVPCDKPFSEEQARLYLRDVLGLYLEYHCKIVHRDIKPSNLL 282

QY 639 I-VDGMLKLIDFGIANQMPTTSVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKSKISP 697

Db 283 LGDDGHVKIADFGVSNQFEGNDAQL--SSTAGTAFMAPEAISD-SGQSFSGKA----- 333

QY 698 KSDVWSLGCILYYMTYKTPFQQIINQISKLHAIDPNHEIEFPDIPF--KDLQDVLKCC 755

Db 334 -LDVWATGVTLYCFVYKCPF--IDDFILALHKKI-KNEPVVFPPEEISEELKDLILKM 389

QY 756 LKRDPKORISIPPELLAHPV 775

Db 390 LDKNPETRIGVDPDKLHPV 409

Search completed: September 29, 2004, 16:17:26

Job time : 39 secs

; ORGANISM: Homo sapiens

US-10-135-689-2

Query Match 6.8%; Score 298.5; DB 4; Length 505;

Best Local Similarity 28.8%; Pred. No. 1.2e-15;

Matches 92; Conservative 58; Mismatches 109; Indels 61; Gaps 14;

QY 493 LASSANECISVKGRISILKQIGSGSKVQVFLNEKK-QIYAIKYNLEADNQ----- 547

Db 114 VAISDAEDCVQLNQ--YKLQSEIGKGGAYGVVRLAYNESEDRHYAMKVLKSKKLLKQYGF 171

QY 548 -----TLDSEYRNEIAYLNKLQ--HSDKIIRLYDYEITDQYIYVM 586

Db 172 RPPPPRGSAAGGPAKQLPLERVYQEIALLKLDHVNVLKIEVLDDPAEDN-LYLVF 230

QY 587 ECGNIDLNSWLKK-----KKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFL 638

Db 231 DL-----LRKGPVMEVPCDKPFSEEQARLYLRDVLGLYLEYHCKIVHRDIKPSNLL 282

QY 639 I-VDGMLKLIDFGIANQMPTTSVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKSKISP 697

Db 283 LGDDGHVKIADFGVSNQFEGNDAQL--SSTAGTAFMAPEAISD-SGQSFSGKA----- 333

QY 698 KSDVWSLGCILYYMTYKTPFQQIINQISKLHAIDPNHEIEFPDIPF--KDLQDVLKCC 755

Db 334 -LDVWATGVTLYCFVYKCPF--IDDFILALHKKI-KNEPVVFPPEEISEELKDLILKM 389

QY 756 LKRDPKORISIPPELLAHPV 775

Db 390 LDKNPETRIGVDPDKLHPV 409

Search completed: September 29, 2004, 16:17:26

Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:11:35 ; Search time 46 Seconds
(without alignments)
1758.632 Million cell updates/sec

Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDIKNKFKNEDLTDEL.....GGESHNSSSSKTFEKRGKK 841

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	4376	100.0	857	2 A42861	protein kinase TTK
2	3182.5	72.7	830	2 B44439	protein kinase (EC
3	3165.5	72.3	856	2 A44439	protein kinase (EC
4	673.5	15.4	678	2 T43539	spindle checkpoint
5	643	14.7	720	2 A96807	hypothetical prote
6	586.5	13.4	764	1 S67561	protein kinase RPK
7	434.5	9.9	536	2 F90132	hypothetical prote
8	326	7.4	460	2 S58882	protein kinase Cds
9	324.5	7.4	576	2 T41587	probable carbon ca
10	321	7.3	587	2 JC5669	Ca2+/calmodulin-de
11	318.5	7.3	1607	2 T03022	MAP kinase kinase
12	317	7.2	1016	2 T25433	hypothetical prote
13	316.5	7.2	417	2 JE0191	calcium/calmodulin
14	309	7.1	1004	2 H88562	protein C07A9.3 [i
15	307.5	7.0	1044	2 S40704	hypothetical prote
16	304.5	7.0	1401	2 T39225	MAP kinase kinase
17	300	6.9	505	2 A57156	Ca2+/calmodulin-de
18	299	6.8	1478	2 S20117	protein kinase BCK
19	297	6.8	746	2 S62365	SNF1-related prote
20	296.5	6.8	357	2 T37317	probable Ca2+/calm
21	296.5	6.8	1097	2 F96538	hypothetical prote
22	295	6.7	1080	2 S48944	hypothetical prote
23	294.5	6.7	515	2 A25048	regulatory protein
24	294.5	6.7	631	2 A57286	probable serine/th
25	294.5	6.7	1081	2 S51899	probable protein k
26	294.5	6.7	1142	2 S50632	protein kinase PAK
27	293.5	6.7	1338	2 T30565	MAP kinase kinase
28	291	6.6	465	2 B55748	protein kinase (EC
29	291	6.6	925	2 A55748	protein kinase (EC

30	290.5	6.6	519	2 C86160	hypothetical prote
31	289.5	6.6	544	2 I49376	p21 activated kina
32	288.5	6.6	610	1 A49082	calcium-dependent
33	288	6.6	648	2 T47988	serine/threonine-p
34	287	6.6	988	1 S35362	protein kinase C (
35	285.5	6.5	1906	1 S68235	myosin-light-chain
36	282.5	6.5	583	2 H84810	probable calcium-d
37	282.5	6.5	887	2 T20941	hypothetical prote
38	281.5	6.4	726	2 T33998	hypothetical prote
39	281	6.4	554	2 T05476	calcium-dependent
40	280.5	6.4	544	2 A57597	beta-p21-activated
41	280.5	6.4	622	2 T15467	hypothetical prote
42	280.5	6.4	639	1 T02784	calcium-dependent
43	280	6.4	560	2 S51600	phosphorylase kina
44	280	6.4	891	2 T40503	protein kinase kin
45	279	6.4	891	2 A38903	protein kinase 1 -

ALIGNMENTS

RESULT 1

A42861
protein kinase TTK (EC 2.7.1.1) - human
N;Alternate names: phosphotyrosine picked threonine kinase (PYT)
C;Species: Homo sapiens (man)
C;Date: 10-Jun-1993 #sequence revision 18-Nov-1994 #text_change 03-Nov-2000
C;Accession: A42861; S27971; I38144
R;Mills, G.B.; Schmandt, R.; McGill, M.; Amendola, A.; Hill, M.; Jacobs, K.; May, C.; R
J. Biol. Chem. 267, 16000-16006, 1992
A;Title: Expression of TTK, a novel human protein kinase, is associated with cell proli
A;Reference number: A42861; MUID:92348472; PMID:1639825
A;Accession: A42861
A;Molecule type: mRNA
A;Residues: 1-857 <MIL1>
A;Cross-references: EMBL:M86699; NID:G340010
A;Note: sequence extracted from NCBI backbone (NCBIN:109875, NCBIP:109876)
A;Note: it is uncertain whether Met-1 or Met-17 is the initiator
R;Mills, G.B.; Schmandt, R.; McGill, M.; Amendola, A.; Hill, M.; Jacobs, K.; May, C.; R
submitted to the EMBL Data Library, February 1992
A;Description: Expression of TTK, a novel human protein kinase.
A;Reference number: S27971
A;Accession: S27971
A;Molecule type: mRNA
A;Residues: 17-857 <MIL2>
A;Cross-references: EMBL:M86699; NID:G340010; PIDN:AAA61239.1; PID:G340011
R;Lindberg, R.A.; Fischer, W.H.; Hunter, T.
Oncogene 8, 351-359, 1993
A;Title: Characterization of a human protein threonine kinase isolated by screening an
A;Reference number: I38144; MUID:93149596; PMID:7678926
A;Accession: I38144
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 'MCMR',504-767,'V',769-802,'GI' <LIN>
A;Cross-references: EMBL:X70500; NID:G312815; PIDN:CAA49912.1; PID:G312816
C;Genetics:
A;Gene: GDB:TTK
A;Cross-references: GDB:455142
A;Map position: 7p12-7cen
C;Superfamily: protein kinase homology
C;Keywords: phosphotransferase
F;523-791/Domain: protein kinase homology <KIN>

Query Match 100.0%; Score 4376; DB 2; Length 857;
Best Local Similarity 100.0%; Pred. NO. 2e-170;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNSGTVNQIMMANNPEDWLSLLKLEKNS 60

Db 17 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNSGTVNQIMMANNPEDWLSLLKLEKNS 76

QY 61 VPLSDALLNKLIGRYSQATBALPPDKYQNESFARIQVFAELKAIQEPDDARDYQFQMAR 120

Db 77 VPLSDALLNKLIGRYSAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR 136

QY 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180

Db 137 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 196

QY 181 EEEKNLSASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 240

Db 197 EEEKNLSASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 256

QY 241 NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKQRTSRSECDLVVPGSKP 300

Db 257 NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKQRTSRSECDLVVPGSKP 316

QY 301 SGNDSCELRNLSKQVNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360

Db 317 SGNDSCELRNLSKQVNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 376

QY 361 PEVPESNQWQAKRKSEKINQNPAASSNHQWQIPELARKVNTQKHTTPEQPVFSVSKQS 420

Db 377 PEVPESNQWQAKRKSEKINQNPAASSNHQWQIPELARKVNTQKHTTPEQPVFSVSKQS 436

QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPPACQLSTPYGQPAQFQQQHH 480

Db 437 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPPACQLSTPYGQPAQFQQQHH 496

QY 481 QILATPLQNLQVLASSANECISVKGRYISILKQISGGSSKVFQVNLNEKKQIYAIKYVN 540

Db 497 QILATPLQNLQVLASSANECISVKGRYISILKQISGGSSKVFQVNLNEKKQIYAIKYVN 556

QY 541 LEEADNQTLDYSYRNEIAYLNKLQHQHSDKIIRLYDYEITDQYIYVMVECGNIDLSWLKXK 600

Db 557 LEEADNQTLDYSYRNEIAYLNKLQHQHSDKIIRLYDYEITDQYIYVMVECGNIDLSWLKXK 616

QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKLKIDFGIANQMOPDIT 660

Db 617 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKLKIDFGIANQMOPDIT 676

QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQ 720

Db 677 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQ 736

QY 721 IINQISKLHAIIDPNHEIEFPDPEKDLQDVLCCLKRDPKQRISEPELLAHYPVQIQTH 780

Db 737 IINQISKLHAIIDPNHEIEFPDPEKDLQDVLCCLKRDPKQRISEPELLAHYPVQIQTH 796

QY 781 PVNQMAKGTTEEMKYVLQVLGLNSPNSILKAAKTLYEYHSGGSHNSSSSKTFEKKRGK 840

Db 797 PVNQMAKGTTEEMKYVLQVLGLNSPNSILKAAKTLYEYHSGGSHNSSSSKTFEKKRGK 856

QY 841 K 841

Db 857 K 857

RESULT 2

B44439

protein kinase (EC 2.7.1.-) esk splice form 2 - mouse

N;Alternate names: probable serine/threonine/tyrosine protein kinase

C;Species: Mus musculus (house mouse)

C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 10-Jul-1998

C;Accession: B44439

R;Douvillle, E.M.; Afar, D.E.; Howell, B.W.; Letwin, K.; Tannock, L.; Ben-David, Y.; Pawl

Mol. Cell. Biol. 12, 2681-2689, 1992

A;Title: Multiple cDNAs encoding the esk kinase predict transmembrane and intracellular

A;Reference number: A44439; MUID:92269841; PMID:1375325

A;Accession: B44439

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-830 <DOU>

A;Experimental source: P19 embryonic carcinoma cell line

A;Note: sequence extracted from NCBI backbone (NCBI:103064, NCBI:103066)

C;Superfamily: protein kinase homology

C;Keywords: phosphotransferase

F;496-764/Domain: protein kinase homology <KIN>

Query Match 72.7%; Score 3182.5; DB 2; Length 830;

Best Local Similarity 74.0%; Pred. No. 4.4e-122;

Matches 622; Conservative 87; Mismatches 105; Indels 27; Gaps 7;

QY 1 MNKVRDIKNFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 60

Db 17 MSKMRDIKNKI-NEDCTDELSLSKICADHTE---TVNQIMRVGNTPENWLNFLKLEKNS 72

QY 61 VPLSDALLNKLIGRYSAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR 120

Db 73 SPLNDDLLNKLIGRYSAIEALPPDKYQNESFARIOVRFAELKAIQEPDDARDYFQMAR 132

QY 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180

Db 133 ENCKKFAFVHVSFAQFELSOGNLKSKSEQLLHKAETGAVFLQMLETAMRNHLQKKQLLP 192

QY 181 EEEKNLSASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 240

Db 193 EEDKKSASTVLSAQEPSSSLGNVQNRISCSERQGAARVLYGENLPPQDAEVRHQ 252

QY 241 NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKQRTSRSECDLVVPGSKP 300

Db 253 NPFKQTHAAKRCSCPFGRVPVNLNSPDPFYVKTDDSSAVTQLTTRKSGPD-RDAILPGSRP 311

QY 301 SGNDSCELRNLSKQVNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360

Db 312 RGSDSYELRGLKPIQTIYLDKSLVSNKSSSEL-MSDLIALKSKTDSS-LTKLEETK---- 365

QY 361 PEVPESNQWQAKRKSEKINQNPAASSNHQWQIPELARKVNTQKHTTPEQPVFSVSKQS 420

Db 366 PEIAERRPMQWQSTRKPECVFQNPAAFAPLRHVPDVTPKAD-----KES 409

QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPPACQLSTPYGQPAQFQQQHH 480

Db 410 PPISVPKWLDPKSACETPSSSSSLDDYMKCFKTPVVKNDPPPACPSSTPYSQLARLQOQQ 469

QY 481 QILATPLQNLQVLASSANECISVKGRYISILKQISGGSSKVFQVNLNEKKQIYAIKYVN 540

Db 470 QGLSTPLQSLQISGSSSINECISVNGRIYSILKQISGGSSKVFQVNLNEKKQINAIKYVN 529

QY 541 LEEADNQTLDYSYRNEIAYLNKLQHQHSDKIIRLYDYEITDQYIYVMVECGNIDLSWLKXK 600

Db 530 LEDADSQTIESYRNEIAFLNKLQHQHSDKIIRLYDYEITEQYIYVMVECGNIDLSWLKXK 589

QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKLKIDFGIANQMOPDIT 660

Db 590 KSINPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKLKIDFGIANQMOPDIT 649

QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYGTPTFQ 720

Db 650 SIVKDSQVGTVMYMAPEAIRDMSSSRENSKIRTKVSPRSDVWSLGCILYMTYGTPTFQ 709

QY 721 IINQISKLHAIIDPNHEIEFPDPEKDLQDVLCCLKRDPKQRISEPELLAHYPVQIQTH 780

Db 710 IINQVSKLHAIINPAHEIEFPPEISEKDLRDLVLCCLVRNPKERISPELLTHPYVQIQPH 769

QY 781 PVNQMAKGTTEEMKYVLQVLGLNSPNSILKAAKTLYEYHSGGSHNSSSSKTFEKKRGK 840

Db 770 PGSQMARGATDEMKYVLQVLGLNSPNSILKTAKTLYERYNCGEGQDSSSSKTFDKKRER 829

841 K 841

830 K 830

RESULT 3

A44439

protein kinase (EC 2.7.1.-) esk splice form 1 - mouse

N;Alternate names: probable serine/threonine/tyrosine protein kinase

C;Species: Mus musculus (house mouse)

[illegible]

Db 293 FDNSQATPIPKRQ-----QDVVTVANLQFIKLVGVGKGGSSVMYRIFSPDNRSLYALK 345
QY 538 YVNLSEADNQLDSYRNEIAYLNKLQHSKDIIRLYDYEITDQ--YIYMMECGNIDLNS 595
Db 346 EVNFINADQTTIGQYKNEIALRLK-SGNDRIIKLYAAEVNDTLGQLNMVMECGETDLAN 404
QY 596 WLKK--KKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKLIDFGIAN 653
Db 405 LLMKNMKPINLNFIRMYWEQMLEAVQVVDQNIHSDLPANFLIVDGMKLIDFGIAK 464
QY 654 QMPDTSVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVNSLGCILYMTY 713
Db 465 AIGNDTTNIHRDISHIGTINYMAPEALTDMAHTNSGVKLVLGRPSDVNSLGCILYQMVY 524
QY 714 GKTFFQIINQISKLHAIIDPNHEIEFPD-----IPEK-----DLQDVLKCC 755
Db 525 GRAPFAH-LKMIQAIAPNEQVHIHFPEVALPANAVQEKESGLPGVTVPDLMDVMKRC 583
QY 756 LKRDPKQRISIPPELLAHPIYVQIQTHFPVNQMAK 787
Db 584 LERDQKRLTIPPELLVHPFLNPLPSYLTPLAK 615
RESULT 5
A96807
hypothetical protein T32E8.5 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A96807
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96807
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-720 <STO>
A;Cross-references: GB:AE005173; NID:g6437532; PIDN:AAF08564.1; GSPDB:GN00141
C;Genetics:
A;Gene: T32E8.5
A;Map position: 1
Query Match 14.7%; Score 643; DB 2; Length 720;
Best Local Similarity 28.1%; Pred. NO. 2.9e-19;
Matches 205; Conservative 99; Mismatches 242; Indels 184; Gaps 24;
QY 107 QEPDDADYFQMARAN-----CKKFAFVHISFAQFELSQGNVKSQQLLQKAVERGAVP 160
Db 30 QRSQDVVTLQSLSAANTLTQDTNNLAITSVAGESASITQPTVSEHFNPSDRQMDFGKSA 89
QY 161 LEMLEIALRNLNKQKQLLSEEEKNLSASTVTAQE-----SPSGSLGHLQNRNNSC-D 214
Db 90 VTSLE---SNLDVQRK-----SQSLIGTSQDMEWDATNOAEASHL---DACIG 131
QY 215 SRGQTKARFLYGENMPPQDAEIGYRNSLRQTNK-TKQSCPFGRVPVNNLNSPCDVKTD 273
Db 132 SKHQ-----NLPSVDSEVSLKSEYKDSLSLAKIQQLGEFP-NFLNQP----- 173
QY 274 DSVVPCFMKQRTSRSECRDLVVPVPGSKPSGNDSCELRNLSVQNSHFKEPLVSEKSEL 333
Db 174 -----RTRC-----SAVGSSWATTLLI 190
QY 334 ITDSITLKNKTESLLAKLEETKEYQEPEVPSNQKQWQAKRKSECINQNPAASSNHQI 393
Db 191 HSSSAPMLNAT-----THVRSRYVEAD-----SNANPHAVQSQGNL 226

QY 394 PELARKVNTQKHTTFEQPVFSVSKSQSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTP 453
Db 227 PSCC-----PSSKVSNIHPNKDATASEMPASTN---DPEVRVKETDTSKQQQITGLEAP 279
QY 454 VVKNDFFPACQLSTPYGPQ-----ACFQQQQHQILATPLQNL---QVLASSANEC----- 501
Db 280 VG-----SSYGSQDQANARLPEELHTSVSSQPKSKDKHEKVASSKGPSAPRKR 329
QY 502 -----ISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIKYVNLSEADNQLDSYRNE 555
Db 330 YDPDLFFKVNKLYQRLKGIGSGSSEVHKVISSDCTIYALKIKLKGRDYATAYGFCQE 389
QY 556 IAYLNKLOQHSDKIIRLYDYEITDQ-----YIYMMECGNIDL 593
Db 390 IGYLKLKGTN-IIQLIDYEVTDKTLLOEVLNGTMSNKGDRVKEDGFYVMVLEYGEIDL 448
QY 594 NSWLKKK-----KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKL 645
Db 449 AHMLSQKWEIEGSDRTIDENWLRFYWQIILQAVNTIHEERIVHSDLPANFLIVRGFLK 508
QY 646 LIDFGIANQMPDTSVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLG 705
Db 509 LIDFGIAKAINSDTTNIQSDSQVGTLSYMSPEAFM-CNESDENGNT-IKCGRPSDIWSLG 566
QY 706 CILYMTYGTPTFPQIINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCLKRDPKQRIS 765
Db 567 CILYQMVYGRTPFADYKTFWAKFKVITDPNHEITYNQLSNPWLIDLMKKCLAWDRNQWR 626
QY 766 IPELLAHPIYV 775
Db 627 IPELLQHPFL 636

RESULT 6

S67561
protein kinase RPK1 (EC 2.7.1.1) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D2785; protein YDL028c; spindle pole body duplication protein
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S67561; S46425; S55256
R;Paulin, L.; Saren, A.M.; Laamanen, P.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67561
A;Accession: S67561
A;Molecule type: DNA
A;Residues: 1-764 <PAU>
A;Cross-references: EMBL:Z74076; NID:g1431003; PIDN:CAA98587.1; PID:g1431004; GSPDB:GN00
A;Experimental source: strain S288C
R;Poch, O.; Schwob, E.; de Fraipont, F.; Camasses, A.; Bordonne, R.; Martin, R.P.
Mol. Gen. Genet. 243, 641-653, 1994
A;Title: RPK1, an essential yeast protein kinase involved in the regulation of the onset
A;Reference number: S46425; MUID:94301294; PMID:8028580
A;Accession: S46425
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-145, 'S', 147-210, 'TKR', 214-764 <POC>
A;Cross-references: GB:L08909; NID:g1197057; PIDN:AAA88731.1; PID:g1197058
R;Lauze, E.; Stoelcker, B.; Luca, F.C.; Weiss, E.; Schutz, A.R.; Winey, M.
EMBO J. 14, 1655-1663, 1995
A;Title: Yeast spindle pole body duplication gene MPS1 encodes an essential dual specific
A;Reference number: S55256; MUID:95255223; PMID:7737118
A;Accession: S55256
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 446-491, 'IERP', 497-715 <LAU>
C;Genetics:
A;Gene: SGD:MPS1; RPK1; MIPS:YDL028C
A;Cross-references: SGD:S0002186; MIPS:YDL028C
A;Map position: 4L
C;Superfamily: yeast protein kinase RPK1; protein kinase homology
C;Keywords: phosphotransferase
F;438-716/Domain: protein kinase homology <KIN>

Db 309 HRDIKPSNLLVGEDGHIKIADFGVSNFEK--GSDALLSNTVGTPTAFMAPES---LSETRK 363

Qy 688 --NGKSKSKI SPKSDVWSLGCILYYMTYKTPFQQIINQISKLHAIIDPNHEIEFPDPIPE 745

Db 364 IFSGKA-----LDVWAMGVTLYCFVGCQPFMD--ERIMCLHSKI-KSQALEFPDQPD 413

Qy 746 --KDLQDLVKCCLKRDPKQORISIPPELLAHPIV 775

Db 414 IAEDLKDLITRMLDKNPESRIVVPEIKLHPWV 445

RESULT 11

T03022

MAP kinase kinase kinase - human

C:Species: Homo sapiens (man)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 05-Nov-1999

C:Accession: T03022

R:Takekawa, M.; Posas, F.; Saito, H.

submitted to the EMBL Data Library, May 1997

A:Description: A human homolog of the yeast Ssk2/Ssk22 MAP kinase kinase kinase, MTK1, A

A:Reference number: Z14824

A:Accession: T03022

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1607 <TAK>

A:Cross-references: EMBL:AF002715; NID:G2352276; PIDN:AAB68804.1; PID:G2352277

C:Genetics:

A:Note: MTK1

Query Match 7.3%; Score 318.5; DB 2; Length 1607;

Best Local Similarity 23.0%; Pred. No. 9.6e-06;

Matches 116; Conservative 98; Mismatches 161; Indels 129; Gaps 23;

Qy 300 PSGNDSCELRNLKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLE----- 353

Db 1206 PSGGDSV---LPKSISSAH-----DTRGSSVPENDRL-----ASIAAELQFRSLR 1248

Qy 354 ---ETKEYQEPEVE-----SNQKQWQAKRKSECINQNPASSNHQIPELARKVNTQK 405

Db 1249 HSSPTEERDEPAYPRGDSGSTRRSWEIR---TLISQSKDTASKLGPIEAIQKSVR--- 1301

Qy 406 HTTFEQPVFSVSKOSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQL 465

Db 1302 --LFEEKRYREMRKNIIG-----QVCDTPKS--YDVMVHVLGRKVT----- 1339

Qy 466 STPYGQPACFQQQHQILATPLQNLQVLASSANECISVKGRISILKQIGSGSSKVFQ 525

Db 1340 -----FKWQRG-----NKIGEGQYCKVYT 1358

Qy 526 VLN-EKKQIYAIVYNLEADNQTLDSYRNEIAYLNKLQHQSDKIIRLYDYEITDQYIYM 584

Db 1359 CISVDGTGELMAMKEIRFPQNDHKTIKETADELKFIFEGIKH--PNLVRYFGVELHREMYI 1416

Qy 585 VME-C--GNIDLSWLKKKSIDPWERKSYWKMLAEAVHTTHQHGIVHSDLKPAN-FLIV 640

Db 1417 FMEYCDGTL EEVSRLGLQEHV---IFLYSKQITIAINVLUHEHGIVHRDIKGANIFLTS 1472

Qy 641 DGMLKLIDFGIANQMPDTSVVKD--SQVGTVNYMPPEAIKDMSSSRENGSKSKISPK 698

Db 1473 SGLIKLGDGFCSVKLXNNAQTMPGEVNSTLGTAAAYMAPEVI-----TRAKGEGHGR---A 1524

Qy 699 SDVWSLGCILYYMTYKTPFQQIINQISKLHAIIDPNHEIEFPDPIPEK---DLQDVLKCC 755

Db 1525 ADIWSLGCVVVIEMVTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKDFLSHC 1580

Qy 756 LKRDPKQORISIPPELLAHPIVQIQT 779

Db 1581 LESDPKMRWTASQLLDHSFVKVCT 1604

RESULT 12

T25433

hypothetical protein ZK524.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T25433; T27888

R:McMurray, A.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z20033

A:Accession: T25433

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1016 <WIL>

A:Cross-references: EMBL:Z72517; PIDN:CAA96698.1; GSPDB:GN00019; CESP:ZK524.4

A:Experimental source: clone T28F4

R:Gardner, A.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z20435

A:Accession: T27888

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1016 <WI2>

A:Cross-references: EMBL:Z73912; PIDN:CAA98148.1; GSPDB:GN00019; CESP:ZK524.4

A:Experimental source: clone ZK524

C:Genetics:

A:Gene: CESP:ZK524.4

A:Map position: 1

A:Introns: 52/3; 91/3; 151/3; 207/1; 254/2; 288/3; 422/2; 560/2; 592/1; 623/1; 660/3; 6

Query Match 7.2%; Score 317; DB 2; Length 1016;

Best Local Similarity 29.2%; Pred. No. 6.7e-06;

Matches 81; Conservative 64; Mismatches 102; Indels 30; Gaps 10;

Qy 508 IYSILKQIGSGSSKVFQVLNEKKQIY-----AIKYVNLEADNQTLDSYRNEIAYLNKL 562

Db 26 LYDLEKTIGQGH---FAVVKLAKHVFTGEMVAVKIIDTKMDEASTSQIMKEVRCM-KL 80

Qy 563 QQHSDKIIRLYDYEITDQYIYVMMECGNIDLSNW-LKKKSIDPWERKSYWKMLAEVHT 621

Db 81 VQHAN-IVRLYEVLDTQTKIFLILELGDYDLHDFIHKHEKGVCSLAQQYFCQIMTAIDY 139

Qy 622 IHQHGIVHSDLKPANFLIVD--GMLKLIDFGIANQMPDTSVVKDSQVGTVNYMPPEAI 679

Db 140 CHQLHVVRDLKPNVNVFEKLGVMVKLTDFGFSNSYEPGEQ---LNTSCGSLAYSAP EIL 196

Qy 680 KDMSSSRENGSKSKISPKSDVWSLGCILYYMTYKTPFQQIINQISKLHAIIDPNHEIE 739

Db 197 -----LGDSYDAPAVDVWSLGVILYMLVCGRLPFQE-ANDSETLT KILDCKYSI- 244

Qy 740 FPDIPKDLQDVLKCOLKRDPKQORISIPPELLAHPIVQ 776

Db 245 -PDVLSDECRNLIQSMLVREPQKRASLEKIVSTSWVQ 280

RESULT 13

JE0191

calcium/calmodulin-dependent protein kinase kinase (EC 2.7.1.-) - human

N:Alternate names: CaMKK

C:Species: Homo sapiens (man)

C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 16-Dec-1998

C:Accession: JE0191

R:Hsu, L.S.; Tsou, A.P.; Chi, C.W.; Lee, C.H.; Chen, J.Y.

Biomed. Sci. 5, 141-149, 1998

A:Title: Cloning, expression and chromosomal localization of human Ca2+/Calmodulin-depe

A:Reference number: JE0191

A:Molecule type: mRNA

A:Residues: 1-417 <HSU>

C:Comment: This enzyme functions as a key intermediaries in Ca2+/calmodulin-driven sign

C:Genetics:

A:Gene: CaMKK

A:Map position: 12

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

C:Keywords: phosphotransferase

F:47-330/Domain: protein kinase homology <KIN>

Query Match	7.2%;	Score 316.5;	DB 2;	Length 417;
Best Local Similarity	25.7%;	Pred. No. 2.7e-06;		
Matches	90;	Conservative	79;	Mismatches 116; Indels 65; Gaps 16;
QY	461	PACQLSTPYGQACFQQ---CQFQILATPLQNQLVLASSANECISVKGRISILKQIGS	517	
Db	11	PYSPVSSPQSSRLPRRPPTVESHVSIQMGQ-----DCVQLNQ--YTLKDEIGK	57	
QY	518	GGSSKVFQVLNEK-KQIYAIKYVNLEADNQT-----LDSY	552	
Db	58	GSYGVVKLAYNENDNTYYAMKVLSSKKLIRAQGPFRPPRGRTRPAPGCGIQRGP	117	
QY	553	RNEIAYLNKLQOQSD--KIIRLYDYEITDQYIYVMVMECGNIDLNSMLKKKSIDPW	610	
Db	118	YQEIALLKKL-DHPNVVLVEVD-DPNEDHLYNVFELVNGQFVMEVPTLKPLSEDQ	175	
QY	611	YWKMLEAVHTIHQHGIVHSDLKPANFLI-VDGMLKLIDFGIANQMQPDTTSVVK	669	
Db	176	YFQDLIKGIEYLYHQKIIHRDIKPSNLLVGEDGHKIADFGVSNEFK--GSDALL	233	
QY	670	TVNYMPEPAIKDMSSRE--NGKSKSKISPKSDVMSLGCILYYMTYKTPFQQIINQ	727	
Db	234	TPAFMAPES---LSETRKIFSGKA-----LDVWAMGVTLYCFVFGQCPFMD--	281	
QY	728	LHAIIDPNHEIEFPDIPE--KDLQDVLKCKLRDPKQKQISIPPELLAHPYV	775	
Db	282	LHSKI-KSQALEFPDQPDIAEDLKDLITRMLDKNPESRIVVPEIKLHPWV	330	
RESULT 14				
H88562	protein C07A9.3 [imported] - Caenorhabditis elegans			
C;Species:	Caenorhabditis elegans			
C;Date:	10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001			
C;Accession:	H88562			
R;anonymous:	The C. elegans Sequencing Consortium.			
Science	282, 2012-2018, 1998			
A;Title:	Genome sequence of the nematode C. elegans: a platform for investigating biolog			
A;Reference number:	A75000; PMID:99069613; PMID:9851916			
A;Note:	see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele			
A;Note:	published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and			
A;Accession:	H88562			
A;Status:	preliminary			
A;Molecule type:	DNA			
A;Residues:	1-1004 <STO>			
A;Cross-references:	GB:chr_III; PIDN:CAA82347.1; PID:g3873989; GSPDB:GN00021; CESP:C07A9			
C;Genetics:				
A;Gene:	C07A9.3			
A;Map position:	3			
C;Superfamily:	unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo			
Query Match	7.1%;	Score 309;	DB 2;	Length 1004;
Best Local Similarity	20.1%;	Pred. No. 1.4e-05;		
Matches	176;	Conservative	159;	Mismatches 326; Indels 214; Gaps 39;
QY	44	NNPEDWLSLLKLEKNSVPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRF-AE	102	
Db	163	SSPTMQCQMAMSEDSIEVRD-----YNSGVHHMHPHQMQMQQQQHHQQQYNMS	213	
QY	103	LKAIQEPDDARDYFQMARANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLE	162	
Db	214	YHNHQOQMQMHYHQOQQYQOQQAQHHQMYAP-QIQQ---QQQQPQQSQSQSAQQPQ	268	
QY	163	MLEIALNINLQKQLLSEEEKNL-SASTVLTAGESFSGSLGHLQNPNNSCDSRGQTTK	221	
Db	269	QSSAALQHVN-----ESSNLSAGSISDRPEQHGCTPQRPTAPQSSTATDKKTR	318	
QY	222	ARFLYGENWPPQDAEIGYRNSLRQTNKTKQSCPFGRVPVNLNLPDCKVKTDDSVVPCFM	281	
Db	319	KRRKAG---PTED-----QATPKQERKITEFMKVGG-----EVASGNSVARCLL	359	
QY	282	K--RQTSRSECRDLVVPKSPGNDSCELR-----NLKSVQNSHF-----	319	

Db	360	TEYHQNGSPKRQ---PAVQNGSNSYDSQQQOPQMNHQMNSYWGVPATPSLGVNNGT	416	
QY	320	-----KEPLVDEKSSSELIITDSITLKNKTESSL-----LAK	351	
Db	417	PTPTQQQHYSSDSNSNSNQSPGQNGSGRMVRI--IDEETQTDSSLSQANPQNADEVAK	474	
QY	352	LEETKEYQEPEVPESNQKQWAKRKSE-----CINQNPAASS----	388	
Db	475	MNRIIEDHRRQIEELNSKNSLERRKNEASKETIKRLIDKNQIERKALRDKTAADSPRIG	534	
QY	389	-----NHQIPELARK---VNTEQKHTTTFEQPVFSVSKQSP-----	421	
Db	535	CFKTTTRTGDSFRDQVDGWAFAPAEKTEQINAEARNEIASASAL--LKKRKLPLGIGKEPK	592	
QY	422	-PISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNQD--FPPACQLSTPY-----	469	
Db	593	RPQAVNSQNDSDNGMQPSTSSNTNGD-DAIFRPEEPKEIQOEYIELDEIYKLRREHLRK	651	
QY	470	-----GQPACFQQQHQILATPLQNQLVLASSANECIS--VKGRIYSILKQIGSGGS	520	
Db	652	EETDLSMEKERLEKEQHVV-----RELKRASNESASQFNDHRLHLKRYLMNLGKGGF	706	
QY	521	SKVFQVLNEKKQIYA---IKYVNLEEADNQTLDSYRNEIAY--LNKLQOQHSDDKIIRLYDY	575	
Db	707	SEVWKAFDIEENRYVACKIHHVKNQWKEEKANYVKHAMREKDIHKSLDHC-RIVKQYDL	765	
QY	576	EITDQVIY-MVMEC--GNIDLNSWLKKKSIDPWERKSYWKNMLEAVHTIHQHG--IVHS	630	
Db	766	LTIDNHSFCTVLEYVPGN-DLDFYLKQNRISSEKARSIIIMQVVSALVYLNEKSTPIIHY	824	
QY	631	DLKPANFLI---VDGMLKLIDFGIANQMOPDT---TSVVKDSQ-VGTVNYMPPEAIKD	681	
Db	825	DLKPANILLESNTSGAIKITDFGLSKIMEGESDDHDLGIELTSQFAGTYWYLPPETFI-	883	
QY	682	MSSSRENGKSKSKISPKSDVMSLGCILYYMTYKTPFQQIINQ--ISKLHAIIDPNHEIE	739	
Db	884	-----VPPPKITCKVDVMSIGVIFYQCIYGGKPGFNDLTQQKILEYNTIINA-REVS	934	
QY	740	FDPDIPE--KDLQDVLKCKLRDPKQKQISIPPELLAH	772	
Db	935	FPSKPQVSSAAQDFIRRCUQYRKEDRADVFELAKH	969	
RESULT 15				
S40704	hypothetical protein C07A9.3 - Caenorhabditis elegans			
C;Species:	Caenorhabditis elegans			
C;Date:	06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 28-May-1999			
C;Accession:	S40704			
R;Smith, M.	submitted to the EMBL Data Library, December 1993			
A;Reference number:	S40701			
A;Accession:	S40704			
A;Status:	preliminary			
A;Molecule type:	DNA			
A;Residues:	1-1044 <SMI>			
A;Cross-references:	EMBL:Z29094; NID:g436440; PID:g436444			
C;Genetics:				
A;Intons:	86/3; 158/3; 265/3; 340/2; 500/3; 980/3; 1017/3			
C;Superfamily:	unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase			
C;Keywords:	ATP; serine/threonine-specific protein kinase			
F;733-1012/Domain:	protein kinase homology <KIN>			
F;741-749/Region:	protein kinase ATP-binding motif			
Query Match	7.0%;	Score 307.5;	DB 2;	Length 1044;
Best Local Similarity	20.0%;	Pred. No. 1.7e-05;		
Matches	176;	Conservative	159;	Mismatches 326; Indels 217; Gaps 39;
QY	44	NNPEDWLSLLKLEKNSVPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRF-AE	102	
Db	200	SSPTMQCQMAMSEDSIEVRD-----YNSGVHHMHPHQMQMQQQQHHQQQYNMS	250	
QY	103	LKAIQEPDDARDYFQMARANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLE	162	

Db 251 YHNHQOQMOMHYHQQOQOYQQOQAQHQMYP-QIQQ-----QQQPPQQOQQOQSAQQPQ 305
Qy 163 MLEIALNLNLQKKQLLSEEEKNL-SASTVLTAQESFSGSLGLQNRNNSCDSRGCTTK 221
Db 306 QSSAALQHVN-----ESSNLSSAGSISDREPEQHGGTQRPPTAQSSSTATDKKTR 355
Qy 222 APFLYGENMPPQDAEIGYRNSLRQTNKTKQSCPPFGRVPVNLNSPDCDVKTDDSVVPCFM 281
Db 356 KRRKAG---PTED-----QATPKQERKITEFMKVGG-----EVASGNSVARCLL 396
Qy 282 K--RQTSRSECRDLVVPGSKPSGNDSCELR-----NLKSVQNSHF----- 319
Db 397 TEYHQNQGSFKRQ---PAVQONGSNSYDSQQQQPQMNHQMHQNSYWG VATPSLGVNNGRT 453
Qy 320 -----KEPLVSDEKSSSELIITDSITLKNKTESSL----- 348
Db 454 PTPTQQQHYSSDSNSNSNQSPPGQGNQSGRMVRT--IDEETQTDSSLSQRFQANQNADE 511
Qy 349 LAKLEETKEYQEPEVPESNQKWQAKRKSE-----CINQNPAASS- 388
Db 512 VAKMNRRIEDHRRQIEELNSKNSLERRKNEASKETIKRLIDKNQIERKALRDKTAADSP 571
Qy 389 -----NHWQIPELARK---VNTEQKHTTPEQPVFSVSKQSP----- 421
Db 572 RIGCFKTRTGTDSFRDQWVDGWAFEMDKTEQINAERNEIASASAL--LKKRKPLGIGK 629
Qy 422 ----PISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKND--PPPACQLSTPY----- 469
Db 630 EPKRPPQAVNSQNDNGMQPSTSSNTNGD-DAIFRRPEEPKEIQYQYIEDELDEIYKLREH 688
Qy 470 -----GQPACFOQQOQHILATPLQNLQVLASSANECIS--VKGRIYSILKQIGS 517
Db 689 LRKEETDLSMEKERLEKEKHV-----RELKRASNESASQFNDHRLLLHKRYLMLNLLGK 743
Qy 518 GGSSKVFQVLNEKKQIYA---IKYVNLEADNQTLDSYRNEIAY--LNKLQOHSDKIURL 572
Db 744 GGFSEVWKAFDIEENRYVACKIHHVKNKWKEKKANYVKHAMPEKDIHKSLDHC-RIVKQ 802
Qy 573 YDYEITDQYIY-MVMEC--GNIDLNSWLKKKSIDPWERKSYKNMLEAVHTIHQHG--I 627
Db 803 YDLLTIDNHSFCTVLEYVPGN-DLDFYLKQNRSISEKEARSIMQVVSALVYLNEKSTPI 861
Qy 628 VHSDLKPANFLI-----VDGMLKLIDFGIANQMQPDT-----TSVVKDSQ-VGTVNYMPPEA 678
Db 862 IHYDLKPANILLESNGNTSGAIKITDFGLSKIMEGESDDHDLGIELTSQFAGTYWYLPPE 921
Qy 679 IKDMSSSRENGKSKSISPKSDVMSLGCILYMYTYGKTPFQIINQ--ISKLHAIDPNH 736
Db 922 FI-----VPPPKITCKVDVWSIGVIFYQCIYGKKPFGNDLTQQKILEYNTIINA-R 971
Qy 737 BIEFPDPIPE--XDLQDVLKCCCLKRDPKORISIPELLAH 772
Db 972 EVSFPSPQVSSAAQDFIRRCLOQRKEDRADVFELAKH 1009

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: September 29, 2004, 16:13:55 ; Search time 138 Seconds
(without alignments)
1922.832 Million cell updates/sec

Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDIKNKFKNEDLTDEL.....GGESHNSSSSKTFEKKRGKK 841

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4373	99.9	857	4 Q9BW51	Q9bw51 homo sapien
2	3092.5	70.7	815	11 Q8BY97	Q8by97 mus musculu
3	2157.5	49.3	882	13 Q98T92	Q98t92 xenopus lae
4	2077.5	47.5	585	11 Q8BX21	Q8bx21 mus musculu
5	1648.5	37.7	983	13 Q8AYG3	Q8ayg3 brachydanio
6	1646	37.6	982	13 Q7T2A4	Q7t2a4 brachydanio
7	1564	35.7	305	4 Q15272	Q15272 homo sapien
8	647.5	14.8	777	10 Q84VX4	Q84vx4 arabidopsis
9	643	14.7	720	10 Q9CA22	Q9ca22 arabidopsis
10	637.5	14.6	742	10 Q8LN55	Q8ln55 oryza sativ
11	637	14.6	630	5 Q9VEH1	Q9veh1 drosophila
12	628	14.4	519	10 P93833	P93833 arabidopsis
13	501.5	11.5	586	5 Q8SSH4	Q8ssh4 encephalito
14	490	11.2	297	10 Q8LA38	Q8la38 arabidopsis
15	434.5	9.9	536	10 Q98S30	Q98s30 guillardia
16	339	7.7	756	13 Q90ZY6	Q90zy6 brachydanio

17	336	7.7	1090	5 Q9GRT3	Q9grt3 leishmania
18	336	7.7	1090	5 Q9N9J2	Q9n9j2 leishmania
19	334	7.6	503	11 Q8BXM8	Q8bxm8 mus musculu
20	334	7.6	541	11 Q8QZT7	Q8qzt7 mus musculu
21	334	7.6	579	11 Q8C0G3	Q8c0g3 mus musculu
22	331	7.6	588	11 Q8C078	Q8c078 mus musculu
23	330	7.5	588	11 Q8CH42	Q8ch42 mus musculu
24	325	7.4	1429	5 Q8IHP1	Q8ihp1 plasmodium
25	324	7.4	465	5 Q8TOL6	Q8tol6 drosophila
26	324	7.4	520	5 Q9VHF6	Q9vhf6 drosophila
27	323.5	7.4	588	4 Q9Y5N2	Q9y5n2 homo sapien
28	322	7.4	588	4 Q86P98	Q86p98 drosophila
29	321	7.3	587	11 Q88831	Q88831 rattus norv
30	320	7.3	2265	5 Q8IID2	Q8iid2 plasmodium
31	318	7.3	1167	5 Q8T2I8	Q8t2i8 dictyosteli
32	317	7.2	557	4 Q94883	Q94883 homo sapien
33	317	7.2	1016	5 Q22855	Q22855 caenorhabdi
34	316.5	7.2	417	4 Q9UER3	Q9uer3 homo sapien
35	316.5	7.2	522	11 Q8K1X6	Q8k1x6 mus musculu
36	316.5	7.2	523	13 Q8AYR1	Q8ayr1 xenopus lae
37	316.5	7.2	533	4 Q96RP2	Q96rp2 homo sapien
38	316.5	7.2	533	4 Q96RR3	Q96rr3 homo sapien
39	316.5	7.2	541	4 Q8WY04	Q8wy04 homo sapien
40	316.5	7.2	541	4 Q8IUG3	Q8iug3 homo sapien
41	316.5	7.2	588	4 Q96RP1	Q96rp1 homo sapien
42	316.5	7.2	588	4 Q96RR4	Q96rr4 homo sapien
43	315	7.2	1497	5 Q95YH6	Q95yh6 drosophila
44	315	7.2	1571	5 Q95YH7	Q95yh7 drosophila
45	315	7.2	1571	5 Q8MSQ4	Q8msq4 drosophila

ALIGNMENTS

RESULT 1

Q9BW51 ID Q9BW51 PRELIMINARY; PRT; 857 AA.
AC Q9BW51; Q9NTM0;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE TTK protein kinase (PPT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye, and Testis;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC000633; AAH00633.1; -
DR EMBL; BC032858; AAH32858.1; -
DR HSSP; Q00534; 1BI8.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 857 AA; 97072 MW; 51F40A3CD1677AC5 CRC64;

Query Match 99.9%; Score 4373; DB 4; Length 857;
Best Local Similarity 99.9%; Pred. No. 1.7e-270;
Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=22354683; PubMed=12466851;
RA	The FANTOM Consortium,
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;
RT	"Analysis of the mouse transcriptome based on functional annotation of
RT	60,770 full-length cDNAs.";
RL	Nature 420:563-573(2002).
DR	EMBL; AK041487; BAC30960.1; -.
DR	MGD; MGI:1194921; Ttk.
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR	InterPro; IPR000719; Prot_kinase.
DR	InterPro; IPR002290; Ser_thr_kinase.
DR	InterPro; IPR008271; Ser_thr_pkin_AS.
DR	InterPro; IPR008941; TPR-like.
DR	InterPro; IPR001245; Tyr_kinase.
DR	Pfam; PF00069; pkinase; 1.
DR	ProDom; PD000001; Prot_kinase; 1.
DR	SMART; SM00220; S_TKC; 1.
DR	SMART; SM00219; TyrKC; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ	SEQUENCE 815 AA; 91524 MW; F12A2C0582E36767 CRC64;
Qy	Query Match 70.7%; Score 3092.5; DB 11; Length 815;
Db	Best Local Similarity 74.2%; Pred. No. 7.9e-189;
Qy	Matches 604; Conservative 83; Mismatches 100; Indels 27; Gaps 7;
Qy	1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNGSGTVNQIMMANNPEDWLSLLKLEKNS 60
Db	17 MSKIRDIKNKI-NEDCTDELSLKICADHTE--TVNQIMRVGNTPENWLNFLKLEKNS 72
Qy	61 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFaelKAIQEPDDDARDYFQMAR 120
Db	73 SPLNDDLNLKLIGRYSQAIEVLPDPKYQNESFARIQVRLAELKAIQEPDDDARDYFQMAR 132
Qy	121 ANCKKFAFVHISFAQFELSQGNVKKSKQLQKAVERGAVPLEMLEIALRNLNKQKQLLS 180
Db	133 ENCKKFAFVHVSFAQFELSQGNLKSEQLLKHAVETGAVPLQMLETAMRNHLQKQLLP 192
Qy	181 EEEKNLKSASTVLTAESEFSGSLGHLQNRNNSCDTSRGQTTTKARFLYGENMPPQDAEIGYR 240
Db	193 EEDKKSASASTVLSAQEPFSSSLGNVQNRISCSERGAARVLYGENLPPQDAEVRHQ 252
Qy	241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVTKDDSVWPCFMKQTSRSECRDLVVPGSKP 300
Db	253 NPFKQTHAAKRSCTPFGRVPVNLNNSPDFYVKTDSSAVTQLTTRKSGSPD-RDAILPGSRP 311
Qy	301 SGNDSCELRNKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESSLLAKLEETKEYQE 360
Db	312 RGSDSYELRGLKPIQTIYLDKSLVNEKSSSEL-MSDLIALKSKTDS-LTKLEETK--- 365
Qy	361 PEVPESNQKQWAKRKSECEINQNPAASSNHWQIPELARKVNTQKHTEFQPVFSVSKQS 420
Db	366 PEIAERRPMQWQSTRKPECVQNPAAFAPLRHVDPVTPKAD-----KES 409
Qy	421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFFPPACQLSTPYGQPACFQQQOH 480
Db	410 PPISVPKWLDPKSACETPSSSSLDYMKCFKTFPVVKNDFFPACPSSTPYSQLARLQOQOQ 469
Qy	481 QILATPLQNLQVLASSANECISVKGRISYLKQIGSGGSSKVQVLINEKKQIYAIFYN 540
Db	470 QGLSTPLQSLQISGSSSINECISVNGRIYSILKQIGSGGSSKVQVLINEKKQINAIFYN 529
Qy	541 LEEADNQTLDSYRNEIAYLNKLCQHSDKIIRLYDYEITDQIYIMWECGNIDLSNLWKKK 600
Db	530 LEDADSQTIESYRNEIAFLNKLQHSDKIIRLYDYEITEQIYIMWECGNIDLSNLWKKK 589

Qy	601	KSIDPWERKSYKWNMLEAVHTTHQHGIVHSDLKPANFLIVDGMKLKIDFGIANQMPDTT	666
Db	590	KSINPWERKSYKWNMLEAVHIIHQHGIVHSDLKPANFIVDGMKLKIDFGIANQMPDTT	649
Qy	661	SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYKTPFQQ	720
Db	650	SIVKDSQVGTVMYMAPEAIRDMSSSRENSKIRTKVSPRSDVMSLGCILYYMTYGRTPFQH	709
Qy	721	IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCCCLKRPDKORISIPELLAHPYVQIQTH	780
Db	710	IINQVSKLHAIINPAHEIEFPEISEKDLRDLVKCCLVNPKERISIPELLTHPYVQIQPH	769
Qy	781	PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAK	814
Db	770	PGSQMARGATDEMKYVLGQLVGLNSPNSILKTAK	803
RESULT 3			
ID	Q98T92	PRELIMINARY;	PRT; 882 AA.
AC	Q98T92;		
DT	01-JUN-2001	(TReMBLrel. 17, Created)	
DT	01-JUN-2001	(TReMBLrel. 17, Last sequence update)	
DT	01-OCT-2003	(TReMBLrel. 25, Last annotation update)	
DE	Mps1/TTK.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxID=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Magnaghi-Jaulin L., Vigneron S., Lorca T., Labbe J.C.;		
RT	"X-Mps1/TTK is a novel Xenopus protein required for the spindle		
RT	checkpoint.";		
RRL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
DR	EMBL; AF250290; AAK27843.1; -.		
DR	HSSP; Q00534; 1B18.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	InterPro; IPR008271; Ser_thr_pkin_AS.		
DR	Pfam; PF00069; pkinase; 1.		
DR	ProDom; PD000001; Prot_kinase; 1.		
DR	SMART; SM00220; S_TKC; 1.		
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.		
DR	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.		
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.		
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.		
SQ	SEQUENCE 882 AA; 99078 MW; 51C0F1B54DFDCD00 CRC64;		
Query Match 49.3%; Score 2157.5; DB 13; Length 882;			
Best Local Similarity 51.7%; Pred. No. 3.9e-129;			
Matches 459; Conservative 134; Mismatches 208; Indels 87; Gaps 22			
Qy	1	MNKVRDIKNKFKNED-LTDELSLNKISADTTDNSGTVNQMIMMANNPEDWLSLLKLEKN	59
Db	17	LDRVKSFKTKYGTDDNWTDELTFKSSADTTEHSGIFTH-LVTTKTPEEWLNLCKLENT	75
Qy	60	SVPLSD-ALLNKLIGRYSQAIEALPPDKYQGNESFARIOVRFAELKAIQEPDARDYFQM	118
Db	76	GLPQIDPQLLNKLIDNYSQAVGALPAEKHSHNESYAKILVHFAELKAIHDLDEAREQFQF	135
Qy	119	ARANCKKFAFVHISFAQFELSQCNGVKKSKOLLQKAVERGAVPLEMLEIALRNLLQKKQL	178
Db	136	ARLNCKKFAFVHTAFAQFELSSEGNFKCKQILRRGLECGAVPCEMLNALKNVQFKKPQL	195
Qy	179	LSEBEKKNL--SASTVLTAQESFSG-SLGHQLQNRNNSCDSRGQTTTKARFLYGENM-PPQD	234

Db	196	ISDEKENVAVSSSHINQGTASFQNLALGNPQ-RMKIESPEEYSVKTRFSYGEKLSPPED	254
QY	235	AE-IGYRNSLRQTNKTKQSCPFGRVPVNNLNSPCDV-KTDDSVVVPCFMKRQTSRSE---	289
Db	255	FEDIGRKPLLMSAKT---CPLGRVPVQPATSPDTRTRKSDGSGSSSVVWRPFSSMRVPV	311
QY	290	-----CRDLVWPGS-----KPSGNDSCELRNLSVQNSHFKEPLVSDSE	327
Db	312	LATLSNPKYLGDDLHCSEDIKVPSSNLPHEEQANEDSLDMKTPSSVILS-VTESNLTVK	370
QY	328	KSSELIITDSITLKNKTESSLLAKLEETK--EYQEPEVPESNQQAQKAKKSECINQNP	385
Db	371	RNEDLIL-----GSTSTAVLNPOENSKPSESAPLLPNIN-----ARLSDTTKPNNI	417
QY	386	ASSNHWQIPELARKVNTQKHTTTE-QPVFSVSKQ-----SPPISTSKWFD	430
Db	418	EAEWKWKVPETPROI-----FQPEMIFPAGKRKSTEPSANPGSRRVSPAPSLSKCD	469
QY	431	PKSICKTPSSNTLDDYMSCFRTPVWKNDFPPACQLSTPYGQACFQQQHQHILATPLQN-	489
Db	470	PVFVCGTPVNKSQEDYMNCFRTPVVKTNLGLAQMSTPYNNRSYQQPQ-----TPVGQP	524
QY	490	--LOVLASSANECISVKGRIYSILKQICSGSGSKVFQVNLNEKKIYAIKYVNLEEDNQ	547
Db	525	DCFPIAPFASDDCIVVKGRAYAVLKQIGTGGSSKVFQVMDDKHLYAIKYVNLEEDQQ	584
QY	548	TLDSYRNEIAYLNKLOQHSDKIIRLYDYEITDQYIYVMMECGNIDLNSWLKKKSIDPWE	607
Db	585	TIESYQNEISHLNKLOQHCCKIIRLFYDYEITDQHIYVMMECGNIDLNTWLRKKKTIINPWE	644
QY	608	RKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLKIDFGIANQMPDPTTSVVKDSQ	667
Db	645	RKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLKIDFGIANQMPDPTTSVVKDSQ	704
QY	668	VGTVNYMPPEAIKDMSSSRENGSKSKISPKSDVMSLGCILYYMTYKTPFQQIINQISK	727
Db	705	VGTINYPPEISIRDTTSYAENGKPRSKISPKGDVMSLGCILYMTYKTPFQHIITNQIAK	764
QY	728	LHAIIDPNHEIEFPDIPEKDLQDVLKCLKRPDKORISIPELLAHPYVQIQTH--PVNQ	785
Db	765	LHSILDPGYEIEFPDIPEKDLQDVLKCLVNRNPKRISIAELLVHPYVQIQPHTQPDQV	824
QY	786	AKGTTEEMKYVLGQLVGLNSPNSILKAAKTLYEHYSGGESHNSSSKT	833
Db	825	QKETTEEMKRILQGLIGLNSPNSISRAAKNLYDQFNSGRSLDLSTLGT	872
RESULT 4			
Q8BX21	PRELIMINARY; PRT; 585 AA.		
ID	Q8BX21		
AC	Q8BX21;		
DT	01-MAR-2003	(TReMBLrel. 23, Created)	
DT	01-MAR-2003	(TReMBLrel. 23, Last sequence update)	
DT	01-OCT-2003	(TReMBLrel. 25, Last annotation update)	
DE	Dual specificity protein kinase TTK (Fragment).		
GN	TTK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs.";		
RL	Nature 420:563-573 (2002).		
DR	EMBL; AK049200; BAC33605.1; --		
DR	MGD; MGI:1194921; Ttk.		
DR	GO; GO:0005524; F:ATP binding; IEA.		

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphotyrosine picked threonine kinase (PRT) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93149596; PubMed=7678926;
RA Lindberg R.A., Fischer W.H., Hunter T.;
RT "Characterization of a human protein threonine kinase isolated by
RT screening an expression library with antibodies to phosphotyrosine."
RL Oncogene 8:351-359(1993).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; X70500; CAA49912.1; -;
DR PIR; A42861; A42861.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 305
SQ SEQUENCE 305 AA; 34971 MW; 2B3F757EEC1CE7D1 CRC64;

Query Match 35.7%; Score 1564; DB 4; Length 305;
Best Local Similarity 99.7%; Pred. No. 6.7e-92;
Matches 298; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 488 QNLQVLASSANECISVKGRISYILKQIGSGSSKVFQVLNKKQIYAIKYVNLEEDNQ 547
Db 5 QNLQVLASSANECISVKGRISYILKQIGSGSSKVFQVLNKKQIYAIKYVNLEEDNQ 64

QY 548 TLDSYRNEIAYLNKLOQHSKIRLYDEYITDQIYVMWMECGNIDLNSWLKKSIDPWE 607
Db 65 TLDSYRNEIAYLNKLOQHSKIRLYDEYITDQIYVMWMECGNIDLNSWLKKSIDPWE 124

QY 608 RKSYYKNNLEAVHTIHHQGVHSDLKIPANFLIVDGMKLIDFGIANQMPDITTSVVKDSQ 667
Db 125 RKSYYKNNLEAVHTIHHQGVHSDLKIPANFLIVDGMKLIDFGIANQMPDITTSVVKDSQ 184

QY 668 VGTVMNPPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYGTPTFQIINQISK 727
Db 185 VGTVMNPPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYGTPTFQIINQISK 244

QY 728 LHAIIDNHEIEFPDIPKIDQVLKCLKRDPKQRIPIPELLAHFYVQIOTHPVNQMA 786
Db 245 LHAIIDNHEIEFPDIPKIDQVLKCLKRDPKQRIPIPELLAHFYVQIOTHPVNQMA 303

RESULT 8
Q84VX4 PRELIMINARY; PRT; 777 AA.
ID Q84VX4
AC Q84VX4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Atlg77720.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones."
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT004636; AA042882.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 777 AA; 86376 MW; F2C25DA9609A878F CRC64;

Query Match 14.8%; Score 647.5; DB 10; Length 777;
Best Local Similarity 27.1%; Pred. No. 7e-33;
Matches 217; Conservative 111; Mismatches 279; Indels 195; Gaps 26;

QY 46 PEDWLSLLKLEKNSVPLSDALLNKLIGRYSQA-----IEALPPDKYQNESFARIQV 98
Db 15 PKSLVRPILNLETTSSSSSSSSPELLRHLQAFKRHRPLSKMQTTSIGPRRSVAPQRQ 74

QY 99 RFAELKAI---QEPDDARDYFQMARAN-----CKKFAFVHISFAQFELSQGNVKKSKQ 148
Db 75 ASRNTRLVTAEGQRSQDVVTLTSQSLAANTLTQDTNNLAITSVAGESASITQPTVSEHFN 134

QY 149 LLQKAVERGAVPLEMLEIALRNINLQKKQLLSEEEKNLSASTVLTAQE-----SFGSL 203
Db 135 PSDRQMDFGKSAVTSLE---SNLDVQRK-----SQSLIGTSQDMEWDATNQAEA 180

QY 204 GHLQNRNNSC-DSRGQTTKARFLYGENMPPQDAEIGYRNSLRQTNK-TKQSCPFGRVPVN 261
Db 181 SHL---DACIGSKHQ-----NLPSVDEVSLKSEYKDSLSLAKIQGLGEFP-N 225

QY 262 LLNSPCDVKTDDSVVPCFMKQRTSRSECRDLVVPGSKPSGNDSCELNLSKVQNSHFE 321
Db 226 FLNQP-----RTRC-----S 235

QY 322 PLVDEKSSSELIITDSITLKNKTESSLLAKLEETKEYQEPEVPESNOKQWAKRKSECIN 381
Db 236 AVGSSWATTTLIHSSSAPMLNAT-----THVSRSYVEAD-----SN 271

QY 382 QNPAASSNHWQIPELARKVNTQKHTTTFEQPVSVSKQSPPISTSKWDFPKSICKTPSSN 441
Db 272 ANPHAVQSQGNLPSCC-----PSSKVSNIHLPNKDATASEMPASTN---DPEVRVKETDTS 324

QY 442 TLDDYMSCFRTPVVKNDFFPACQLSTPYGQP-----ACFQQQHQILATPLQNL---QVLA 494
Db 325 KQQQITGTLEAPVG-----SSIYGSDGQANARLPEELHTSVSSQPKSKDKHEKVA 374

QY 495 SSSANEC-----ISVKGRISYILKQICSGGSSKVFQVLNKKQIYAIKYVNLEE 543
Db 375 SSKGPSAPKRKNYDPPDLFFKVGKLYQRLGKIGSGSSSEVHKVISSDCTIYALKKILKG 434

QY 544 ADNQTLDSYRNEIAYLNKLOQHSKIRLYDEYITDQ-----Y 581
Db 435 RDTATAYGFCQEIYGLYKLLKGGKTN-IIQLIDYEVTDKTLQEVNLNGTMSNKDGRVKEDGF 493

QY 582 IYVMWMECGNIDLNSWLKKS-----KSIDPWERKSYKNNLEAVHTIHHQGVHSDLK 633
Db 582 IYVMWMECGNIDLNSWLKKS-----KSIDPWERKSYKNNLEAVHTIHHQGVHSDLK 633

D b	494	IYMVLEYGEIDLAAHMLSQKWREIEGSDRTIDENWLRFYWQOILQAVNTIHEERIVHSDLK	553
QY	634	PANFLIVDGMKLKIDFGIANQMOPDTTSVVKDSQVGTVNVNPPAEAIKDMSSSRENGKSKS	693
D b	554	PANFLLVRGFLKLIDFGIAKAINSDTNIQRDSQVGTLSYNPFEAFM-CNESDENGNT-I	611
QY	694	KISPKSDVWSLGCILYYMTYGTPTFQOIINQISKLHAIIDPNHIEIEFPDIPEKDQDVLK	753
D b	612	KCGRPDSIWSLGCILYQMVYGRTPFADYKTFWAKFVITDPNHEITYNQLSNPWLIDLMK	671
QY	754	CCLKRDPKQORISIPELLAHPYV	775
D b	672	KCLAWDRNQWRIPELLQHPFL	693

RESULT 9
Q9CA22 PRELIMINARY; PRT; 720 AA.
ID Q9CA22
AC Q9CA22;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T32E8.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.
RA Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.-L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Ienz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC012193; AAG51619.1; --
DR PIR; A96807; A96807.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 720 AA; 80027 MW; F26A9A0B1C6D91DE CRC64;

Best Local Similarity 28.1%; Pred. No. 1.2e-32;
Matches 205; Conservative 99; Mismatches 242; Indels 184; Gaps 24;

```

QY 107 QEPDDARDYQMARAN-----CKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVP 160
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30 QRSQDVVTLSSLAANTLTQDTTNNLAITSVAGESASITQPTVSEHFNPSRQMDFGKSA 89
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 LEMLEIALRNILCKQLLSEEEKKNLSASTVLTAE-----SFGSLGHLQNRNNSC-D 214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 VTSLE---SNLDVQRK-----SQSLIGTSQDMEWDATNQAESH-----DACIG 131
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 215 SRGQTTKARPLYGENMPPQDAEIGYRNSLRQTNK-TKQSCPFGRVPVNLNLSPPDCDVKT 273
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 SKHQ-----NLPVSDSEVSLKSEYKSSSLAKIQGLGEFP-NFLNQ-----173
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 274 DSVVPCFMKQTSRSECRDLVVPGPSKSGNDSCELRNLKSVQNSHFKEPLVDSKSELI 333
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174 -----RTRC-----SAVGSSWATTTLI 190
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 334 ITDSITLKNKTESSLLAKLEETKEYQEPEVPESNQWQAKRKSECINQNPAASSNHQI 393
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 HSSAPMLNAT-----THVRSYVEAD-----SNANPHAVQSQGNL 226
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 394 PELARKVNTQKHTTPEQPVFSVSKQSPPISTSKWEDPKSICKTPSSNTLDDYMSCFRTP 453
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 PSCC-----PSSKVSNIHPNKDATASENPASTN---DPEVRVKETDTSKQQQITTGLEAP 279
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 454 VVKNDFFPACQLSTPYGQP---ACFQQQHQILATPLNL---QVLASSANEC-----501
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 VG-----SSYIGSDGQANARLPEELHTSVSSQPKSKDKHEKVASSKGPSAPRKRN 329
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 502 -----ISVKGRISILKQIGGGSSKVQVLNEKKQIYAIKYVNLEEDNQTLDSYRNE 555
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 YDPDLFFKVNGLYQRLGKIGSGSSEVHKVITSSDCTIYALKKIKLGRDYATAYGFCQE 389
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 IAYLNKLQOHSDKIIRLYDYEITDQ-----YIYVMVECGNIDL 593
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 IGYLXKLKGXTN-IIQLIDYEVTDKTLQEVLTMSNKGDRVKEDGFIYMLVLEYGEIDL 448
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 594 NSWLXXX-----KSIDPWERKSYWKNNLEAVHTIHQHGIVHSDLKPANFLIVDGMKL 645
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 AHMLSQKWRBIEGSDRTIDENWLRFYWQQLQAVNTIHEERI VHSDLKPANFLIVRGFLK 508
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 646 LIDFGIANQWQPDTSVVKDSQVGTVNYMPPEAIKDMSSSRENGKSKSIKSPKSDVWSLG 705
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 LIDFGIAKAINSDTTNIQDSQVGTLSYMSPEAFM-CNESDENGNT- IKGRPSDIWSLG 566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 CILYYMTYKTPFQIINQISKLHAIIDPNHEIEFPDPEKDLQDVLCCKLRDPKQRIS 765
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567 CILYQMVYGRTPFADYKTFWAKFKVITDPNHEITYNQLSNPWLIDLMKKCLAWDRNQWR 626
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 766 IPELLAHPYV 775
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 627 IPELLQHPFL 636
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10
Q8LN55
ID Q8LN55 PRELIMINARY; PRT; 742 AA.
AC Q8LN55;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

RESULT 10		
Q8LN55		
ID	Q8LN55	PRELIMINARY;
AC	Q8LN55;	PRT; 742 AA.
DT	01-OCT-2002	(TREMBLrel. 22, Created)
DT	01-OCT-2002	(TREMBLrel. 22, Last sequence update)
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)
DE	Hypothetical protein.	

OS *Oryza sativa* (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; *Oryza*.

ID Q8SSH4 PRELIMINARY; PRT; 586 AA.
AC Q8SSH4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MP51-like THR/TYR dual specificity protein kinase.
GN EC02_0510.
OS Encephalitozoon cuniculi.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL; AL590442; CAD25082.1; --
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Transferase.
KW ATP-binding; Transferase.
SQ SEQUENCE 586 AA; 66975 MW; CBE4BA154FBBB6A CRC64;
Query Match 11.5%; Score 501.5; DB 5; Length 586;
Best Local Similarity 41.7%; Pred. No. 9.9e-24;
Matches 116; Conservative 51; Mismatches 96; Indels 15; Gaps 7;
QY 498 ANECISVKGRISILKQIGSGSSKVFQVLNKKQIYAIKYVNLEEDNQTLDYSRNEIA 557
Db SRERIIKGREIEILKQIGKSSKVKYKVL-FGSNVYALKRVEL-IGDEKMLSSYINEIN 335
QY 558 YLNLKQHSQDKIIRLYDYETDQYIYVMVECGNIDLSWLKKKSIDPWERKSYKKNMLE 617
Db LLYKFKGTSE-IVEIIDHEVGEDYLHILLEYGETDL-SKIIRKGGLSMNFIDVWEQMLL 393
QY 618 AVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTTSVVKDSQVGTVNYMPPE 677
Db IVKRVHTIERIIHCDLKPANFLVKGVRVKLIDFGISKVIRNDTTSILSBEQCCTVNYMSP 453
QY 678 AIKDMSSSRENGKSKSKISPKSDVWSLGCILYVNTYGTPTFQIINQISKLHAIIDPNHE 737
Db 454 AVT-----QNKSKVARSSDIWSLGCILYVMVHSNPPLHYPNLIQIQLQEYS-E 503
QY 738 IEFDPDPEKLDQVLKCKLCKRDPKQRISEIPPELLAHPYV 775
Db 504 FKYTS-KNKAAMVMKECLARDPKKRPTIDNLLNHRFL 540
RESULT 14
Q8LA38 PRELIMINARY; PRT; 297 AA.
AC Q8LA38;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY088049; AAM65595.1; --
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Transferase.
SQ SEQUENCE 297 AA; 34366 MW; A07227F3044D5E0 CRC64;
Query Match 11.2%; Score 490; DB 10; Length 297;
Best Local Similarity 50.2%; Pred. No. 2.1e-23;
Matches 103; Conservative 29; Mismatches 63; Indels 10; Gaps 3;
QY 579 DQYIYVMVECGNIDLSWLKKK-----KSIDPWERKSYKKNMLEAVHTIHQHGIVHS 630
Db 11 DGFYIMVLEYGEIDLAHMLSQKWEIEGSDRTIDENWLFYQQILQAVNTIHEERIVHS 70
QY 631 DLKPANFLIVDGMKLIDFGIANQMOPDTTSVVKDSQVGTVNYMPPEAIKDMSSSRENGK 690
Db 71 DLKPANFLIVRGFLKLIDFGIAKAINSDTTNIDQDSQVGTLSYMSPEAFM-CNESDENG 129
QY 691 SKSKTSPKSDVWSLGCILYVNTYGTPTFQIINQISKLHAIIDPNHEIEFPDPEKLDQ 750
Db 130 T-IKGRPSDIWSLGCILYQMVYGRTPFADYKTFWAKFKVITDPNHEITYNQLSNPW 188
QY 751 VLKCKLCKRDPKQRISEIPPELLAHPYV 775
Db 189 LMKKCLAWDRNQRWRIPPELLQHPFL 213
RESULT 15
Q98S30 PRELIMINARY; PRT; 536 AA.
ID Q98S30
AC Q98S30;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein kin from chromosome 3.
GN KIN.
OS Guillardia theta (Cryptomonas phi).
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223349; PubMed=11323671;
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: September 29, 2004, 16:16:51 ; Search time 144 Seconds
(without alignments)
1879.397 Million cell updates/sec

Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDIKNKFKNEDLTDEL.....GGESHNSSSSKTFEKKRGKK 841

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4376	100.0	841	14	US-10-081-119-14
2	4376	100.0	841	14	US-10-116-712-666
3	4376	100.0	841	15	US-10-353-690-2
4	4376	100.0	841	15	US-10-366-288-16
5	4376	100.0	841	15	US-10-295-027-804
6	4376	100.0	841	15	US-10-295-027-833
7	4376	100.0	841	15	US-10-295-027-865
8	4376	100.0	857	14	US-10-354-358-30
9	4373	99.9	857	15	US-10-295-027-482
10	4373	99.9	857	15	US-10-173-999-34
11	1929	44.1	367	16	US-10-664-421-157
12	673.5	15.4	678	14	US-10-081-119-32
13	637.5	14.6	742	16	US-10-437-963-201272
14	627	14.3	730	12	US-10-424-599-273093
15	614	14.0	399	12	US-10-425-114-43102

16	603	13.8	735	14	US-10-128-714-3040	Sequence 3040, Ap
17	603	13.8	842	14	US-10-128-714-8040	Sequence 8040, Ap
18	586.5	13.4	764	14	US-10-081-119-30	Sequence 30, Appl
19	586.5	13.4	764	15	US-10-369-493-1595	Sequence 1595, Ap
20	558	12.8	690	14	US-10-032-585-7510	Sequence 7510, Ap
21	553.5	12.6	399	15	US-10-369-493-3420	Sequence 3420, Ap
22	352.5	8.1	256	12	US-09-964-956-72	Sequence 72, Appl
23	352.5	8.1	256	12	US-10-042-865-83	Sequence 83, Appl
24	352.5	8.1	256	12	US-10-029-020-183	Sequence 183, App
25	352.5	8.1	256	12	US-10-072-012-799	Sequence 799, App
26	352.5	8.1	256	12	US-10-072-012-856	Sequence 856, App
27	352.5	8.1	256	12	US-10-072-012-876	Sequence 876, App
28	352.5	8.1	256	12	US-10-072-012-882	Sequence 882, App
29	352.5	8.1	256	16	US-10-055-569A-99	Sequence 99, Appl
30	349.5	8.0	254	15	US-10-074-978A-150	Sequence 150, App
31	345.5	7.9	255	12	US-10-087-684-98	Sequence 98, Appl
32	345.5	7.9	255	12	US-10-218-779-98	Sequence 98, Appl
33	344.5	7.9	256	10	US-09-823-187-96	Sequence 96, Appl
34	343.5	7.8	256	10	US-09-863-776-40	Sequence 40, Appl
35	342.5	7.8	749	9	US-09-771-161A-261	Sequence 261, App
36	340.5	7.8	251	15	US-10-052-648A-41	Sequence 41, Appl
37	338	7.7	528	14	US-10-032-585-7571	Sequence 7571, Ap
38	335.5	7.7	252	15	US-10-074-978A-160	Sequence 160, App
39	332.5	7.6	250	10	US-09-976-782-31	Sequence 31, Appl
40	329.5	7.5	749	12	US-10-092-900A-20	Sequence 20, Appl
41	328.5	7.5	250	12	US-10-042-865-102	Sequence 102, App
42	328.5	7.5	787	14	US-10-354-358-48	Sequence 48, Appl
43	328.5	7.5	787	16	US-10-408-765A-267	Sequence 267, App
44	327	7.5	460	9	US-09-740-627-10	Sequence 10, Appl
45	319	7.3	231	9	US-09-515-806-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-081-119-14
; Sequence 14, Application US/10081119
; Publication No. US20030045491A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Jefferson, Anne B.
; APPLICANT: Chan, Vivien W.
; TITLE OF INVENTION: TTK in Diagnosis and as a Therapeutic
; TITLE OF INVENTION: Target in Cancer
; FILE REFERENCE: 16932.002
; CURRENT APPLICATION NUMBER: US/10/081,119
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/289,813
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 14
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-081-119-14

Query Match		100.0%;	Score 4376;	DB 14;	Length 841;
Best Local Similarity		100.0%;	Pred. No. 2.4e-299;		
Matches		841;	Conservative	0;	Mismatches
				0;	Indels
				0;	Gaps
				0;	Gaps
QY	1	MNKVRDIKNKFKNEDLTDEL	SLNKISADTTD	NSGT	VNQIMMANNP
Db	1	MNKVRDIKNKFKNEDLTDEL	SLNKISADTTD	NSGT	VNQIMMANNP
QY	61	VPLSDALLNKLI	GRYSQA	TEALPP	DKYQNESF
Db	61	VPLSDALLNKLI	GRYSQA	TEALPP	DKYQNESF
QY	121	ANCKKFAFVHIS	FAQFELS	QGNVKS	QKLLQK
Db	121	ANCKKFAFVHIS	FAQFELS	QGNVKS	QKLLQK

QY	181	EEKNLSASTVLTAESEFSGSLGHLQNRNNSCDVRGQTTKARFLYGENMPPQDAEIGYR	240
Db	181	EEKNLSASTVLTAESEFSGSLGHLQNRNNSCDVRGQTTKARFLYGENMPPQDAEIGYR	240
QY	241	NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP	300
Db	241	NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP	300
QY	301	SGNDSCELRNLKSVQNSHFKEPLVDSKSELITDSTLKNKTESLLAKLEETKEYQE	360
Db	301	SGNDSCELRNLKSVQNSHFKEPLVDSKSELITDSTLKNKTESLLAKLEETKEYQE	360
QY	361	PEVPESNQWQAKRKSECINQNPAASSNHQIPELARKVNTQKHTTPEQPVFSVSKQS	420
Db	361	PEVPESNQWQAKRKSECINQNPAASSNHQIPELARKVNTQKHTTPEQPVFSVSKQS	420
QY	421	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPPACQLSTPYGQACFQQQHH	480
Db	421	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPPACQLSTPYGQACFQQQHH	480
QY	481	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAICYVN	540
Db	481	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAICYVN	540
QY	541	LEEADNQTLDYSYRNEIAYLNKLOQHSDKIIRLYDYEITDQIYIMVMECGNIDLSWLKXK	600
Db	541	LEEADNQTLDYSYRNEIAYLNKLOQHSDKIIRLYDYEITDQIYIMVMECGNIDLSWLKXK	600
QY	601	KSIDPWERKSYWKQMLEAVHTIHQHGIVHSDLPANFLIVDGMKLIDFGIANQMPDTH	660
Db	601	KSIDPWERKSYWKQMLEAVHTIHQHGIVHSDLPANFLIVDGMKLIDFGIANQMPDTH	660
QY	661	SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQQ	720
Db	661	SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQQ	720
QY	721	IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLRDPKQRIISIPELLAHPYVQIQTH	780
Db	721	IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLRDPKQRIISIPELLAHPYVQIQTH	780
QY	781	PVNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYEHYSGGSHSSSSKTFEKKRGK	840
Db	781	PVNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYEHYSGGSHSSSSKTFEKKRGK	840
QY	841	K 841	
Db	841	K 841	

RESULT 2
US-10-116-712-666
; Sequence 666, Application US/10116712
; Publication No. US20030194764A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Switzer, Ann
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.568
; CURRENT APPLICATION NUMBER: US/10/116,712
; CURRENT FILING DATE: 2002-04-07
; NUMBER OF SEQ ID NOS: 670
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 666
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-712-666

Query Match 100.0%; Score 4376; DB 14; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNKVRDIKNKPKNEDLTDELSLNKISADTTDNGSTVNOIMMANNPEDWLSLLKLEKNS	60
Db	1	MNKVRDIKNKPKNEDLTDELSLNKISADTTDNGSTVNOIMMANNPEDWLSLLKLEKNS	60
QY	61	VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFAELKAIQEPDDARDYFQMAR	120
Db	61	VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFAELKAIQEPDDARDYFQMAR	120
QY	121	ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALRNLNLQKQLLS	180
Db	121	ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALRNLNLQKQLLS	180
QY	181	EEKNLSASTVLTAESEFSGSLGHLQNRNNSCDVRGQTTKARFLYGENMPPQDAEIGYR	240
Db	181	EEKNLSASTVLTAESEFSGSLGHLQNRNNSCDVRGQTTKARFLYGENMPPQDAEIGYR	240
QY	241	NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP	300
Db	241	NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP	300
QY	301	SGNDSCELRNLKSVQNSHFKEPLVDSKSELITDSTLKNKTESLLAKLEETKEYQE	360
Db	301	SGNDSCELRNLKSVQNSHFKEPLVDSKSELITDSTLKNKTESLLAKLEETKEYQE	360
QY	361	PEVPESNQWQAKRKSECINQNPAASSNHQIPELARKVNTQKHTTPEQPVFSVSKQS	420
Db	361	PEVPESNQWQAKRKSECINQNPAASSNHQIPELARKVNTQKHTTPEQPVFSVSKQS	420
QY	421	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPPACQLSTPYGQACFQQQHH	480
Db	421	PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDPPPACQLSTPYGQACFQQQHH	480
QY	481	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAICYVN	540
Db	481	QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQIYAICYVN	540
QY	541	LEEADNQTLDYSYRNEIAYLNKLOQHSDKIIRLYDYEITDQIYIMVMECGNIDLSWLKXK	600
Db	541	LEEADNQTLDYSYRNEIAYLNKLOQHSDKIIRLYDYEITDQIYIMVMECGNIDLSWLKXK	600
QY	601	KSIDPWERKSYWKQMLEAVHTIHQHGIVHSDLPANFLIVDGMKLIDFGIANQMPDTH	660
Db	601	KSIDPWERKSYWKQMLEAVHTIHQHGIVHSDLPANFLIVDGMKLIDFGIANQMPDTH	660
QY	661	SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQQ	720
Db	661	SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYMTYKTPFQQ	720
QY	721	IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLRDPKQRIISIPELLAHPYVQIQTH	780
Db	721	IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLRDPKQRIISIPELLAHPYVQIQTH	780
QY	781	PVNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYEHYSGGSHSSSSKTFEKKRGK	840
Db	781	PVNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYEHYSGGSHSSSSKTFEKKRGK	840
QY	841	K 841	
Db	841	K 841	

RESULT 3
US-10-353-690-2
; Sequence 2, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary

APPLICANT: Stagliano, Nancy
APPLICANT: Perodin, Jacqueline
APPLICANT: Rodrigue-Way, Amelie
TITLE OF INVENTION: Methods and compositions for treating
TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
FILE REFERENCE: MPI02-018P1RNMNIM
CURRENT APPLICATION NUMBER: US/10/353,690
CURRENT FILING DATE: 2003-01-29
PRIOR APPLICATION NUMBER: 60/353,224
PRIOR FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 60/364,529
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: 60/373,861
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/376,287
PRIOR FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 60/388,080
PRIOR FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: 60/390,971
PRIOR FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/394,130
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: 60/394,797
PRIOR FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: 60/404,904
PRIOR FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: 60/405,450
PRIOR FILING DATE: 2002-08-23
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 126
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 841
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-353-690-2

Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 60
Db 1 MNKVRDIKNFKNEDLTDELSLNKISADTTDNGTGNQIMMANNPEDWLSLLKLEKNS 60
QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR 120
Db 61 VPLSDALLNKLIGRYSQAIEALPPDKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR 120
QY 121 ANCKKFAFVHISFAQFELSOGNVKKSQILLQKAVERGAVPLEMLEIALRNLNLQKKQLLS 180
Db 121 ANCKKFAFVHISFAQFELSOGNVKKSQILLQKAVERGAVPLEMLEIALRNLNLQKKQLLS 180
QY 181 EEKKNLSASTVLTAAQESFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
Db 181 EEKKNLSASTVLTAAQESFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
QY 241 NSLRQTNKTKQSCPFGRVPVNLNLLNSPDCDKVTDSDVVPFCMKRQTSRSECRDLVVPGSKP 300
Db 241 NSLRQTNKTKQSCPFGRVPVNLNLLNSPDCDKVTDSDVVPFCMKRQTSRSECRDLVVPGSKP 300
QY 301 SGNDSCELRLNLSQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 301 SGNDSCELRLNLSQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
QY 361 PEVPESNQWQAKRKSECINQNPAASNHWQIPELARKVNTQKHTTTEQPVFSVSKQS 420

Db 361 PEVPESNQWQAKRKSECINQNPAASNHWQIPELARKVNTQKHTTTEQPVFSVSKQS 420
QY 421 PPSTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPPACQLSTPYGQPACFQQQOH 480
Db 421 PPSTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPPACQLSTPYGQPACFQQQOH 480
QY 481 QILATPLONLOVLASSANECISVKGRISYILKQIGSGGSKVFQVLNEKKQIYAIFYN 540
Db 481 QILATPLONLOVLASSANECISVKGRISYILKQIGSGGSKVFQVLNEKKQIYAIFYN 540
QY 541 LEEADNQTLDYSRNEIAYLNKLQOHSDKIIRLYDYEITDQYIYVMECGNIDLSWLKKK 600
Db 541 LEEADNQTLDYSRNEIAYLNKLQOHSDKIIRLYDYEITDQYIYVMECGNIDLSWLKKK 600
QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT 660
Db 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT 660
QY 661 SVVKDSQVGTVNYMPPEAKDMSSSRRENGSKSKISPKSDVWSLGCILYYMTYKTPFQ 720
Db 661 SVVKDSQVGTVNYMPPEAKDMSSSRRENGSKSKISPKSDVWSLGCILYYMTYKTPFQ 720
QY 721 IINQISKLHAIIDPNHEIEFPDIPEKLOQVLCCLKRDPKQRISEPELLAHYVQIQTH 780
Db 721 IINQISKLHAIIDPNHEIEFPDIPEKLOQVLCCLKRDPKQRISEPELLAHYVQIQTH 780
QY 781 PVNQMAKGTTEEMKYVLQVLGNLSPNSILKAAKTLIYHSGGESHNSSSSKTFEKRGK 840
Db 781 PVNQMAKGTTEEMKYVLQVLGNLSPNSILKAAKTLIYHSGGESHNSSSSKTFEKRGK 840
QY 841 K 841
Db 841 K 841

RESULT 4

US-10-366-288-16
Sequence 16, Application US/10366288
Publication No. US20030216288A1
GENERAL INFORMATION:
APPLICANT: Powell, Douglas
APPLICANT: Weich, Nadine S.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: AIDS AND HIV-RELATED DISORDERS USING 1414, 1481, 1553,
TITLE OF INVENTION: 34021, 1720, 1683, 1552, 1682, 1675, 12825, 9952, 5816,
TITLE OF INVENTION: 10002, 1611, 1371, 14324, 126, 270, 312, 167, 326, 18926,
TITLE OF INVENTION: 6747, 1793, 1784 OR 2045 MOLECULES
FILE REFERENCE: MPI02-025P1RNMNIM
CURRENT APPLICATION NUMBER: US/10/366,288
CURRENT FILING DATE: 2003-02-13
PRIOR APPLICATION NUMBER: 60/357,391
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 60/380,249
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: 60/391,306
PRIOR FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: 60/406,297
PRIOR FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: 60/412,007
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: 60/417,508
PRIOR FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: 60/432,318
PRIOR FILING DATE: 2002-12-10
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 841
TYPE: PRT
ORGANISM: Homo Sapien
US-10-366-288-16

Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNGSVNQIMMANNPEDWLSLLKLEKNS 60
DB 1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNGSVNQIMMANNPEDWLSLLKLEKNS 60

QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQONESFARIQVFAELKAIQEPDDARDYFQMAR 120
DB 61 VPLSDALLNKLIGRYSQAIEALPPDKYQONESFARIQVFAELKAIQEPDDARDYFQMAR 120

QY 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180
DB 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180

QY 181 EEEKNLASASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENNPPQDAEIGYR 240
DB 181 EEEKNLASASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENNPPQDAEIGYR 240

QY 241 NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP 300
DB 241 NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP 300

QY 301 SGNDSCELRNKLSVQNSHFKEPLVSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
DB 301 SGNDSCELRNKLSVQNSHFKEPLVSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360

QY 361 PEVPESNQWQAKRKSEKINQPAASNHWQIPELARKVNTQKHTTPEQPVFSVSKQS 420
DB 361 PEVPESNQWQAKRKSEKINQPAASNHWQIPELARKVNTQKHTTPEQPVFSVSKQS 420

QY 421 PPISTKWFDPKSIKTPSSNTLDDYMSCFRTPVVKNDFFPPACQLSTPYGQPACFQQQKH 480
DB 421 PPISTKWFDPKSIKTPSSNTLDDYMSCFRTPVVKNDFFPPACQLSTPYGQPACFQQQKH 480

QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQISGGSSKVFQVLEKQIYAIKYVN 540
DB 481 QILATPLQNLQVLASSANECISVKGRIYSILKQISGGSSKVFQVLEKQIYAIKYVN 540

QY 541 LEEADNQTLDSYRNEIAYLNKLQOHSDKIIRLYDYBITDQYIYVMMECGNIDLSWLKXK 600
DB 541 LEEADNQTLDSYRNEIAYLNKLQOHSDKIIRLYDYBITDQYIYVMMECGNIDLSWLKXK 600

QY 601 KSIDPWERKSYWKMLAEVHTTHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT 660
DB 601 KSIDPWERKSYWKMLAEVHTTHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTT 660

QY 661 SVVKDSQVGTVNYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYKTPFQQ 720
DB 661 SVVKDSQVGTVNYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYKTPFQQ 720

QY 721 IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCLKRDPKQKORISIPELLAHYPVQIOTH 780
DB 721 IINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCLKRDPKQKORISIPELLAHYPVQIOTH 780

QY 781 PVNQMAKGTTEEMKYVLGVLNSPNSILKAAKTYEHYSGGESHSNSSSKTTFEKKRGK 840
DB 781 PVNQMAKGTTEEMKYVLGVLNSPNSILKAAKTYEHYSGGESHSNSSSKTTFEKKRGK 840

QY 841 K 841
DB 841 K 841

RESULT 5

US-10-295-027-804
; Sequence 804, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 804
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-804

Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNGSVNQIMMANNPEDWLSLLKLEKNS 60
DB 1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNGSVNQIMMANNPEDWLSLLKLEKNS 60

QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQONESFARIQVFAELKAIQEPDDARDYFQMAR 120
DB 61 VPLSDALLNKLIGRYSQAIEALPPDKYQONESFARIQVFAELKAIQEPDDARDYFQMAR 120

QY 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180
DB 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180

QY 181 EEEKNLASASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENNPPQDAEIGYR 240
DB 181 EEEKNLASASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENNPPQDAEIGYR 240

QY 241 NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP 300
DB 241 NSLRQTNKTKQSCPFGRVPVNLNSPDCDVKTDDSVVPCFMKRQTSRSECRDLVVPGSKP 300

QY 301 SGNDSCELRNKLSVQNSHFKEPLVSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
DB 301 SGNDSCELRNKLSVQNSHFKEPLVSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360

QY 361 PEVPESNQWQAKRKSEKINQPAASNHWQIPELARKVNTQKHTTPEQPVFSVSKQS 420
DB 361 PEVPESNQWQAKRKSEKINQPAASNHWQIPELARKVNTQKHTTPEQPVFSVSKQS 420

QY 421 PPSTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPPACOLSTPYGQPACFOQOQH 480
Db |||||
QY 481 QILATPLQNLQVLASSANECISVKGRISYILKQIGSGGSSKVFQVLNEKKQIYAICYVN 540
Db |||||
QY 541 LEEADNQTLDYSYRNEIAYLNKLOQHSCKIIRLYDYEITDQIYVMVECGNIDLSWLKKK 600
Db |||||
QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTH 660
Db |||||
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYGTPTFQ 720
Db |||||
QY 721 IINQISKLHAIIDPNHEIEFPDPIPEKLDQVLKCCCLKRDPKQRISEPELLAHPYVQIOTH 780
Db |||||
QY 781 PVNQMAKGTTEEMKYVLGQVLGNLSPNSILKAAKTLYEYHSGGSHNSSSKTFEKKRGK 840
Db |||||
QY 841 K 841
Db 841 K 841

RESULT 6
US-10-295-027-833
; Sequence 833, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 833
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-833

Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNFKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 60
Db |||||
QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQONESFARIQVRFAELKAIQEPDDARDYFQMAR 120
Db |||||
QY 121 ANCKKFAFVHISFAQFELSOGNVKSKQLQKAVERGAVPLEMLEIALRNLNLOKKQLLS 180
Db |||||
QY 181 EEEKKNLSASTVLTAESEFSGSLGHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR 240
Db |||||
QY 241 NSLRQTNKTKQSCPFGRVPVNLNLPDCDVKTDDSVVPCFMKQRTSRSECRDLVPGSKP 300
Db |||||
QY 301 SGNDSCELRNLSKVSQNSHFKPLVSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db |||||
QY 361 PEVPESNQKQWAKRKSEGINQNPAASSNHWQIPELARKVNTQKHTTFEQPVFSVSKQS 420
Db |||||
QY 421 PPSTSKWFDPKSICKTPSSNTLDDYMSCRTPVVKNDFFPPACOLSTPYGQPACFOQOQH 480
Db |||||
QY 481 QILATPLQNLQVLASSANECISVKGRISYILKQIGSGGSSKVFQVLNEKKQIYAICYVN 540
Db |||||
QY 541 LEEADNQTLDYSYRNEIAYLNKLOQHSCKIIRLYDYEITDQIYVMVECGNIDLSWLKKK 600
Db |||||
QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLIDFGIANQMOPDTH 660
Db |||||
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYGTPTFQ 720
Db |||||
QY 721 IINQISKLHAIIDPNHEIEFPDPIPEKLDQVLKCCCLKRDPKQRISEPELLAHPYVQIOTH 780
Db |||||
QY 781 PVNQMAKGTTEEMKYVLGQVLGNLSPNSILKAAKTLYEYHSGGSHNSSSKTFEKKRGK 840
Db |||||
841 K 841
841 K 841

RESULT 7
US-10-295-027-865
; Sequence 865, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 865
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-865
Query Match 100.0%; Score 4376; DB 15; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.4e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNKVRDIKNFKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 60
Db 1 MNKVRDIKNFKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 60
QY 61 VPLSDALLNKILGRYSQAIEALPPDKYQNESFARIQVRFAELKAIQEPDARDYFQMAR 120
Db 61 VPLSDALLNKILGRYSQAIEALPPDKYQNESFARIQVRFAELKAIQEPDARDYFQMAR 120
QY 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALENLNQKKLLS 180
Db 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALENLNQKKLLS 180
QY 181 EEEKNLASVTLTAQESFSGSLCHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
Db 181 EEEKNLASVTLTAQESFSGSLCHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
QY 241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 300
Db 241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 300

QY 301 SGNDSCLELNKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESSLAKLEETKEYQE 360
Db 301 SGNDSCLELNKSVQNSHFKEPLVSDKSSSELIITDSITLKNKTESSLAKLEETKEYQE 360
QY 361 PEVPESNOKQWQAKRKSECCINQNPAASSNHWQIPELARKVNTQKHTTTEQPVFVSQKS 420
Db 361 PEVPESNOKQWQAKRKSECCINQNPAASSNHWQIPELARKVNTQKHTTTEQPVFVSQKS 420
QY 421 PPISTSKWFPDKSICKTPSSNTLDDYMSCFRTPVVKNDFFPACQLSTPYGQPACFQQQOH 480
Db 421 PPISTSKWFPDKSICKTPSSNTLDDYMSCFRTPVVKNDFFPACQLSTPYGQPACFQQQOH 480
QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKICGSGSSKVFQVLNEKKQIYAIKYVN 540
Db 481 QILATPLQNLQVLASSANECISVKGRIYSILKICGSGSSKVFQVLNEKKQIYAIKYVN 540
QY 541 LEEADNQTLDSYRNEIAYLNKLQOHSDKIIIRLYDYEITDQIYIYMMECNGIDLSWLKKK 600
Db 541 LEEADNQTLDSYRNEIAYLNKLQOHSDKIIIRLYDYEITDQIYIYMMECNGIDLSWLKKK 600
QY 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLKIDFGIANQMOPDTT 660
Db 601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIVDGMKLKIDFGIANQMOPDTT 660
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSSRENGKSKISPKSDVWSLGCILYIMTYGKTPFQQ 720
Db 661 SVVKDSQVGTVMYMPPEAIKDMSSSSRENGKSKISPKSDVWSLGCILYIMTYGKTPFQQ 720
QY 721 IINQISKLHAIIDPNHEIEFPDIPEKLDQVLKCLKRDPKQRISIPPELLAHPYVQIQTH 780
Db 721 IINQISKLHAIIDPNHEIEFPDIPEKLDQVLKCLKRDPKQRISIPPELLAHPYVQIQTH 780
QY 781 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAATLYEHYSGGESHNSSSKTFFKRGK 840
Db 781 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAATLYEHYSGGESHNSSSKTFFKRGK 840
QY 841 K 841
Db 841 K 841
RESULT 8
US-10-354-358-30
; Sequence 30, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MF102-020P1RNOVIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507

```

; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 857
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-354-358-30

Query Match      100.0%; Score 4376; DB 14; Length 857;
Best Local Similarity 100.0%; Pred. No. 2.5e-299;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNKVRDIKNFKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 60
Db      17 MNKVRDIKNFKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 76

Qy      61 VPLSDALLNKLIGRYSAIEALPPDPKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR 120
Db      77 VPLSDALLNKLIGRYSAIEALPPDPKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR 136

Qy      121 ANCKKFAFVHISFAQFELSQGNVKKSKOLLQKAVERGAVPLEMLEIALNLNLQKQLLS 180
Db      137 ANCKKFAFVHISFAQFELSQGNVKKSKOLLQKAVERGAVPLEMLEIALNLNLQKQLLS 196

Qy      181 EEEKNLASSTVLTAGEFSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 240
Db      197 EEEKNLASSTVLTAGEFSFSGSLGHLQNRNNSCDSRGQTTKARFLYGENMPPQDAEIGYR 256

Qy      241 NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDDSVVPCFMKROTISRSECRDLVFGSKP 300
Db      257 NSLRQTNKTKQSCPFGRVPVNLNNSPCDVKTDDSVVPCFMKROTISRSECRDLVFGSKP 316

Qy      301 SGNDSCELRNLKSVQNSHFKEPLVSDEKSELITDSTITLKNKTESLLAKLEETKEYQE 360
Db      317 SGNDSCELRNLKSVQNSHFKEPLVSDEKSELITDSTITLKNKTESLLAKLEETKEYQE 376

Qy      361 PEVPESNQWQAKRKSECINQNPAASSNHQWQIPELARKVNTQKHITTFEQPVFVSQKS 420
Db      377 PEVPESNQWQAKRKSECINQNPAASSNHQWQIPELARKVNTQKHITTFEQPVFVSQKS 436

Qy      421 PPISTSKWFDPKSICKTPSSNTLDDYMSCERTPVVKNDFPPACQLSTPYGQPACFQQQOH 480
Db      437 PPISTSKWFDPKSICKTPSSNTLDDYMSCERTPVVKNDFPPACQLSTPYGQPACFQQQOH 496

Qy      481 QILATPLQNLQVLASSANECISVKGRIVSILKQIGSGGSKVPQVLNEKKQIYAIKYVN 540
Db      497 QILATPLQNLQVLASSANECISVKGRIVSILKQIGSGGSKVPQVLNEKKQIYAIKYVN 556

Qy      541 LEEADNQTLDYSYRNEIAYLNKLQHQHSDKIIRLYDYEITDQYIYMWVECGNIDLSNLKXK 600
Db      557 LEEADNQTLDYSYRNEIAYLNKLQHQHSDKIIRLYDYEITDQYIYMWVECGNIDLSNLKXK 616

Qy      601 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKLIDFGIANQMOPDPT 660
Db      617 KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIVDGMKLIDFGIANQMOPDPT 676

Qy      661 SVVKDSQVGTVNYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYGTPTFFQ 720
Db      677 SVVKDSQVGTVNYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTYGTPTFFQ 736
```

```

Qy      721 IINQISKLHAIIDPNHEIEFPDPIPEKDQLQDVLKCCCLKRDPKQIRISIPELLAHPYVQIQTH 780
Db      737 IINQISKLHAIIDPNHEIEFPDPIPEKDQLQDVLKCCCLKRDPKQIRISIPELLAHPYVQIQTH 796

Qy      781 PVNQMAKGTTEEMKYVLGQVLGNLSPNSILKAAKTLYEYHSGGESHNSSSKTFEKRGK 840
Db      797 PVNQMAKGTTEEMKYVLGQVLGNLSPNSILKAAKTLYEYHSGGESHNSSSKTFEKRGK 856

Qy      841 K 841
Db      857 K 857
```

RESULT 9

```

US-10-295-027-482
; Sequence 482, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 482
; LENGTH: 857
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-482
```

```

Query Match      99.9%; Score 4373; DB 15; Length 857;
Best Local Similarity 99.9%; Pred. No. 4e-299;
Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 MNKVRDIKNFKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 60
Db      17 MNKVRDIKNFKNEDLTDELSLNKISADTTNSGTVNQIMMANNPEDWLSLLKLEKNS 76

Qy      61 VPLSDALLNKLIGRYSAIEALPPDPKYGNESFARIOVRFAELKAIQEPDDARDYFQMAR 120
```

Db 77 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFAELKAIQEPDDARDYFQMAR 136
QY 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180
Db 137 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 196
QY 181 EEEKKNLSASTVLTAEQSFSGSLGHLQNRNNSCDSDRGQTTKARFLYGENMPPQDAEIGYR 240
Db 197 EEEKKNLSASTVLTAEQSFSGSLGHLQNRNNSCDSDRGQTTKARFLYGENMPPQDAEIGYR 256
QY 241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 300
Db 257 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 316
QY 301 SGNDSCELRNLSKVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 317 SGNDSCELRNLSKVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 376
QY 361 PEVPESNQWQAKRKSECCINQNPAASSNHQIPELARKVNTQKHTTPEQPVFVSQSKS 420
Db 377 PEVPESNQWQAKRKSECCINQNPAASSNHQIPELARKVNTQKHTTPEQPVFVSQSKS 436
QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQQQKH 480
Db 437 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQQQKH 496
QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLENEKKQIYAIFYN 540
Db 497 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLENEKKQIYAIFYN 556
QY 541 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYEITDQYIYVMMECGNIDLSWLKXK 600
Db 557 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYEITDQYIYVMMECGNIDLSWLKXK 616
QY 601 KSIDPWERKSYWKNMLEAVHTIHQGHIVHSDLPANFLIVDGMKLIDFGIANQMOPDTT 660
Db 617 KSIDPWERKSYWKNMLEAVHTIHQGHIVHSDLPANFLIVDGMKLIDFGIANQMOPDTT 676
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYGTPTFQ 720
Db 677 SVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYGTPTFQ 736
QY 721 IINQISKLHAIIDPNHEIEFPDIPEKDLQDLVKCCLKRDPKQISIPPELLAHYVQIOTH 780
Db 737 IINQISKLHAIIDPNHEIEFPDIPEKDLQDLVKCCLKRDPKQISIPPELLAHYVQIOTH 796
QY 781 PVNQMAKGTTEEMKYVLQVLGNSPNSILKAAKTLIYEHYSGGESHNSSSKTFEKKRGK 840
Db 797 PVNQMAKGTTEEMKYVLQVLGNSPNSILKAAKTLIYEHYSGGESHNSSSKTFEKKRGK 856
841 K 841
857 K 857

RESULT 10
US-10-173-999-34
; Sequence 34, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27

; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver: 2.1
; SEQ ID NO 34
; LENGTH: 857
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-34
Query Match 99.9%; Score 4373; DB 15; Length 857;
Best Local Similarity 99.9%; Pred. No. 4e-299;
Matches 840; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNSGTNVQIMMANNPEDWLSLLKLEKNS 60
Db 17 MNKVRDIKNKFKNEDLTDELSLNKISADTTDNSGTNVQIMMANNPEDWLSLLKLEKNS 76
QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFAELKAIQEPDDARDYFQMAR 120
Db 77 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFAELKAIQEPDDARDYFQMAR 136
QY 121 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 180
Db 137 ANCKKFAFVHISFAQFELSOGNVKSKQLLQKAVERGAVPLEMLEIALRNINLQKKQLLS 196
QY 181 EEEKKNLSASTVLTAEQSFSGSLGHLQNRNNSCDSDRGQTTKARFLYGENMPPQDAEIGYR 240
Db 197 EEEKKNLSASTVLTAEQSFSGSLGHLQNRNNSCDSDRGQTTKARFLYGENMPPQDAEIGYR 256
QY 241 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 300
Db 257 NSLRQTNKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECRDLVVPGSKP 316
QY 301 SGNDSCELRNLSKVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
Db 317 SGNDSCELRNLSKVQNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 376
QY 361 PEVPESNQWQAKRKSECCINQNPAASSNHQIPELARKVNTQKHTTPEQPVFVSQSKS 420
Db 377 PEVPESNQWQAKRKSECCINQNPAASSNHQIPELARKVNTQKHTTPEQPVFVSQSKS 436
QY 421 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQQQKH 480
Db 437 PPISTSKWFDPKSICKTPSSNTLDDYMSCFRTPVVKNDFPPACQLSTPYGQACFQQQKH 496
QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLENEKKQIYAIFYN 540
Db 497 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSKVFQVLENEKKQIYAIFYN 556
QY 541 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYEITDQYIYVMMECGNIDLSWLKXK 600
Db 557 LEEADNQTLDYSYRNEIAYLNKLQOHSDKIIRLYDYEITDQYIYVMMECGNIDLSWLKXK 616
QY 601 KSIDPWERKSYWKNMLEAVHTIHQGHIVHSDLPANFLIVDGMKLIDFGIANQMOPDTT 660
Db 617 KSIDPWERKSYWKNMLEAVHTIHQGHIVHSDLPANFLIVDGMKLIDFGIANQMOPDTT 676
QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYGTPTFQ 720
Db 677 SVVKDSQVGTVMYMPPEAIKDMSSSRENGKSKISPKSDVWSLGCILYMTYGTPTFQ 736
QY 721 IINQISKLHAIIDPNHEIEFPDIPEKDLQDLVKCCLKRDPKQISIPPELLAHYVQIOTH 780
Db 737 IINQISKLHAIIDPNHEIEFPDIPEKDLQDLVKCCLKRDPKQISIPPELLAHYVQIOTH 796
QY 781 PVNQMAKGTTEEMKYVLQVLGNSPNSILKAAKTLIYEHYSGGESHNSSSKTFEKKRGK 840
Db 797 PVNQMAKGTTEEMKYVLQVLGNSPNSILKAAKTLIYEHYSGGESHNSSSKTFEKKRGK 856
841 K 841

Db 857 K 857

RESULT 11

US-10-664-421-157

; Sequence 157, Application US/10664421

; Publication No. US20040142864A1

; GENERAL INFORMATION:

; APPLICANT: BREMER, RYAN

; APPLICANT: IBRAHIM, PRABHA

; APPLICANT: KUMAR, ABHINAV

; APPLICANT: MANDIYAN, VALSAN

; APPLICANT: MILBURN, MICHAEL V.

; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE

; FILE REFERENCE: 039363/0703

; CURRENT APPLICATION NUMBER: US/10/664,421

; CURRENT FILING DATE: 2003-09-16

; PRIOR APPLICATION NUMBER: 60/412,341

; PRIOR FILING DATE: 2002-09-20

; PRIOR APPLICATION NUMBER: 60/411,398

; PRIOR FILING DATE: 2002-09-16

; NUMBER OF SEQ ID NOS: 169

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 157

; LENGTH: 367

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-664-421-157

Query Match 44.1%; Score 1929; DB 16; Length 367;

Best Local Similarity 100.0%; Pred. No. 1.7e-127;

Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 FPPACQLSTPYGQPACFQQQHQIILATPLQNLQVLASSANECISVKGRIYSILKQIGSG 518

Db 1 FPPACQLSTPYGQPACFQQQHQIILATPLQNLQVLASSANECISVKGRIYSILKQIGSG 60

QY 519 GSSKVFQVLNEKKQIYAIKYVNLEEADNQTLDSYRNEIAYLNKLOQHSCKIIRLYDYEIT 578

Db 61 GSSKVFQVLNEKKQIYAIKYVNLEEADNQTLDSYRNEIAYLNKLOQHSCKIIRLYDYEIT 120

QY 579 DQYIYMMECGNIDLNSWLKKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFL 638

Db 121 DQYIYMMECGNIDLNSWLKKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFL 180

QY 639 IVDGMLKLIDFGIANQMPDTSVVKDSQVGTVNYMPPPEAIKDMSSSRENGSKSKISPK 698

Db 181 IVDGMLKLIDFGIANQMPDTSVVKDSQVGTVNYMPPPEAIKDMSSSRENGSKSKISPK 240

QY 699 SDWWSLGCILYYMTYKTPFQIINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLKR 758

Db 241 SDWWSLGCILYYMTYKTPFQIINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLKR 300

QY 759 DPKQRISIPPELLAHPYVQIQTHPPVNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYE 818

Db 301 DPKQRISIPPELLAHPYVQIQTHPPVNQMAKGTTEEMKYVLGQVLGNSPNSILKAAKTLYE 360

QY 819 HYSGGES 825

Db 361 HYSGGES 367

RESULT 12

US-10-081-119-32

; Sequence 32, Application US/10081119

; Publication No. US20030045491A1

; GENERAL INFORMATION:

; APPLICANT: Reinhard, Christoph

; APPLICANT: Jefferson, Anne B.

; APPLICANT: Chan, Vivien W.

; TITLE OF INVENTION: TTK in Diagnosis and as a Therapeutic

; TITLE OF INVENTION: Target in Cancer

; FILE REFERENCE: 16932.002

; CURRENT APPLICATION NUMBER: US/10/081,119

; CURRENT FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: 60/289,813

; PRIOR FILING DATE: 2001-02-21

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 32

; LENGTH: 678

; TYPE: PRT

; ORGANISM: Schizosaccharomyces pombe

US-10-081-119-32

Query Match 15.4%; Score 673.5; DB 14; Length 678;

Best Local Similarity 30.9%; Pred. No. 1.1e-38;

Matches 195; Conservative 102; Mismatches 210; Indels 125; Gaps 23;

QY 233 QDAEIGYRNSLRQTNKTKQSCPFGRVPVNLNLSPCDVKTDSDVVPFCFWKQTSRSECRD 292

Db 32 QDPELYFKN---DTFSSKSS-----HSDGTVTGDTLRRQSSGATALE 70

QY 293 LVV--PGSK-----PSGNDSCELR--NLKSVQNSHFKEPLVSDKSSSELIITDSITLKNK 343

Db 71 RLVSHPRTKNFDLQNGGQNSALKKEVNTPAYQSMHFEHLITPLPS-----TNASH 121

QY 344 TESSLLAKLEETKEYOEPE-VPESNQK-----QWQAKRKSECINQNPAASSNHWQIPELA 397

Db 122 SEVSLSAGVNDLNSNSEHDLPLPKSVNKTPGSLSISRRRRIGRIGLGP-----PKRA 172

QY 398 R-----KVNTQKHTTTEQPVFSVSKQSPPISTS---KWFDPKSICTPSSNTLDDVM 447

Db 173 EYTLTDPKSTSDTKNSTEADIEDIMKSREVSPASNSVAATTLKPLQLHNTPLQTSQEHKP 232

QY 448 SCFRTPVVKNDFPPACQL-----STPYQOP-ACFQQQ-----HQI----- 482

Db 233 PSFHPQSQFESSFPRVQFDHVERRASSELHSPVTVFQEPQRSASQPYESHALSPKVAPL 292

QY 483 ----LATPLQNLQVLASSANECISVKGRIYSILKQIGSGSKVQVLN-EKKQIYAIK 537

Db 293 FDNQATPIPKRQ-----QDVVTVANLQFIKLGWVGKGGSMVYRIFSPDNSRLYALK 345

QY 538 YVNLEEADNQTLDSYRNEIAYLNKLOQHSCKIIRLYDYEITDQ--YIYMMECGNIDLNS 595

Db 346 EVNFINADQTTIQGYKNEIALRLKL-SGNDRIIKLYAAEVNDTLGQLNMVMECETDLAN 404

QY 596 WLKK--KKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLPANFLIIVDGMLKLIDFGIAN 653

Db 405 LLMKNMKPINLNFIRYWEQMLEAVQVVDQNVHSDLPANFLIVEGNLKLIDFGIAK 464

QY 654 QMQPDTTSVVKDSQVGTVNYMPPPEAIKDMSSSRENGSKSKISPKSDVWSLGCILYYMTY 713

Db 465 AIGNDTTNIHRDSDHIGTINYMAPEALTDMAHTNSGVKLVKLGRPSDVWSLGCILYYMVY 524

QY 714 GXTFPQQIINQISKLHAIIDPNHEIEFPD-----IPEK-----DLQDVLKCC 755

Db 525 GRAFFAH-LKMIQAIAPNEQYHIHFPEVALPANAVQEKESLPGVTGPDLMVDMKRC 583

QY 756 LKRDPPQRISIPPELLAHPYVQIQTHPPVNQMAK 787

Db 584 LERDQKRLLTIPELLVHFFLNPLPSYLTPLAK 615

RESULT 13

US-10-437-963-201272

; Sequence 201272, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 201272
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_96664C.1.pap
US-10-437-963-201272

Query Match 14.6%; Score 637.5; DB 16; Length 742;
Best Local Similarity 30.6%; Pred. No. 4.2e-36;
Matches 181; Conservative 92; Mismatches 179; Indels 139; Gaps 19;
QY 308 LRNLKSVQNSHFKEPLVSDKSSSLIITDSITL-----KNKTESSLLA-K 351
Db 118 LRNAAPDQ---KAVVSSQ--DELLLTPLTLGTTTTHDQNGQNHQPKSDTDLVDR 171
QY 352 LEETKEYQEPPEVPSNQKQWAKRK-----SECINQNPA-----SSNHWQI 393
Db 172 KKSSMEVSSSQMASANALVGEDFKDLFLYLTSDPQLTSQTDGNFPVAQVADDQGNHKEI 231
QY 394 PELARKVNTQKHTTFFQPVFVSQSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTP 453
Db 232 GIASAAVENDIKY-----DAANLSRRIDEACDQNHGEPWTRCSAMGSSV--TAVSLYSGP 284
QY 454 VVKNDFPFAC-QLSTPYGPACFQ-----QQHQI----- 482
Db 285 TIQSKSAQIDQYASPAQMPQCGRESSGVSGHGSQKLHGVMNHADCNKQOVDNNGM 344
QY 483 -----LATPLQNL-----QVLA-----SSANEC-----ISVKGRIYSI 511
Db 345 DKPVSSSAVCLPSQGLSGNDQSLSAKDDGAPRRSRKVEKCRKKNYDPDFKVGKLYQK 404
QY 512 LKQIGSGSSKVFQVLNEKKQIYAIKYVNLEEAQNTLDSYRNEIAYLNKLQOHSKIIIR 571
Db 405 LGKIGSGSSSEVHKVISAECIIYALKIKLKGRDYPTAYGFCQEIYLYNKLKGSN-IIQ 463
QY 572 LYDYEITDQ-----YIYVMECGNIDL-----NSWLKKKS---ID 604
Db 464 LIDYEVTDKSLQDDSLSPRGRKIDDDHYIYVLEGEIDLANNVAQEWKGRNTSNMKID 523
QY 605 PWERKSYVKNMLEAVHTIHQHGIVHSDLKPANFLVDGMLKLIDFGIANQMOPDTTSVVK 664
Db 524 ENWLFYWOQMLKAVNTIHEERIVHSDLKPANFLVRGALKLIDFGIAKAIMNDTTNIQR 583
QY 665 DSQVGTVMYMPPEAIKOMSSRENGKSKSISPKSDVWSLGCILYMYTYGKTPFQOIINQ 724
Db 584 DSQIGTLNYSPEAF--MCNEQDSGGNVIKCGRPSDIWSLGCILYQVYGYKTPFADYKNF 641
QY 725 ISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLRDPKQISIPPELLAHPYV 775
Db 642 WAKFKVYTDKNHKIKYEPVDNPNWLDLMQRCCLAWDRNRWRIPQLLEHPFL 692

RESULT 14
US-10-424-599-273093
; Sequence 273093, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273093
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(730)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_88625C.1.pap
US-10-424-599-273093

Query Match 14.3%; Score 627; DB 12; Length 730;
Best Local Similarity 29.4%; Pred. No. 2.3e-35;
Matches 192; Conservative 98; Mismatches 243; Indels 120; Gaps 19;
QY 215 SRGQTTKARFLYGEN---MPPQDAEIGYRNSLRQTNKTKQSCFPGRVPVNLNS--PDC 268
Db 86 SVGAHNKVQFLTEINNATAQPDQDEASIGFNQHEKHQNMQQAESLTVSEGGKTSMLPKR 145
QY 269 DVKTDDSV--VPCFMKQTSRSECRDLVVPGSKESGNDSCELRNLSVQN-----SHF- 319
Db 146 TIVTQDHLQOQFKNFLRQPAQSSVVGLPCPTTTSVHSTAPMLNSITHNSCIDSGSHVA 205
QY 320 KEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQEPEVPEPNQKQWAKRKSEC 379
Db 206 AEPYGNLDGNSRPILQGDVKNP-----ISLKDNRMSIDQVASAVQ-----DC 249
QY 380 INQNPAASNHWQIPELARKVNTQKHTTFFQPVFVSQSPPISTSKWFDPKSICKTPS 439
Db 250 --NSP-----IDAEVTVKQSDPSYEQQGCMKDKTISKCTS-----CR--- 285
QY 440 SNTLDDYMSCFRTPVVKNDPFPACQLSTPYGPACFQQQHQIILATPLQNLQVLASSAN 499
Db 286 ----DDMLSKGEVSAATNLQPVPTSS-----SEVKLESSKLEKQEKXPSSKAS 331
QY 500 ECI-----SVKGRISYILKQIGSGSSKVFQVLNEKKQIYAIKYVNLEEAQNT 548
Db 332 SGLRKRAYDPPELFFKVGNGKLYQRLGKIGSGSSSEVHKVVISDCTIYALKRIKLGRDYAT 391
QY 549 LDSYRNEIAYLNKLQOHSKIIIRLYDYEITDQ-----YIYVMV 586
Db 392 AYGFQCEIEYLNRL-KGKNIIQLIDYEVTDKALPEGVINGSFNSKNGRVKDDGIYIYVVL 450
QY 587 ECGNIDLNSWLKK-----KSIDPWERKSYVKNMLEAVHTIHQHGIVHSDLKPANFL 638
Db 451 EYGEIDLHMLSQKWKELDSNQTIDENWLFYWOQILQAVNTIHEERIVHSDLKPANFL 510
QY 639 IVDGMLKLIDFGIANQMOPDTTSVVKDSQVGTVMYMPPEAIKOMSSRENGKSKSISPK 698
Db 511 LVKGSLLKIDFGIAKAIMSDTTNIQDSQVGTLSYMSPEAF--MCNESDASGNIKCGRP 568
QY 699 SDVWSLGCILYMYTYGKTPFQOIINQISKLHAIIDPNHEIEFPDIPEKDLQDVLKCKLR 758
Db 569 SDIWSLGCILYQVYGYRTFADYKTFWAKFKVITDPNHEIMYAPVSNPWLDDLMRRCLAW 628
QY 759 DPKQRISIPPELLAHPYVQIQTHPVNQMAKGTTEEMKYVLGQLVG---LNSPNS 808
Db 629 DRNERWRIPQLLQHPFLV-----PPVPCSHSSFYQEQTFLQLLQISEACINDPES 677

RESULT 15
US-10-425-114-43102
; Sequence 43102, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43102
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700729819_FLI.pep
US-10-425-114-43102

Query Match	14.0%	Score 614;	DB 12;	Length 399;
Best Local Similarity	41.1%	Pred. No. 8.1e-35;		
Matches 139;	Conservative 53;	Mismatches 106;	Indels 40;	Gaps 6;

Qy	504	VKGRIYSILKQIGSGSSKVFQVLNEKKQIYAIFYVNLEADNQTLDYSRNEIAYLNKLQ	563
Db	16	VNGKLYQRLGKIGSGSSEVHKVISSDCTIYALKRIKLGKRDYATAYGFCQEIEYLNRL-	74
Qy	564	QHSDKIIRLYDYETDQ-----YIYMWMECGNIDLSWLKKK-	600
Db	75	KGKNNIIQLIDYEVTDKALFEGVINGSFNKDGVRVKDDGYIYMWLEYGEIDLAHMLSQKW	134
Qy	601	-----KSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKFPANFLIVDGMKLIDFGIAN	653
Db	135	KELDGSNQTIDENWLRFYWQQILQAVNTIHEERIVHSDLKFPANFLLVKGSCLKLIDFGIAK	194
Qy	654	QMOPDTTSVVKDSQVGTVNYMPPEAIKDMSSSRENGKSKSIKSPKSDVWSLGCILYMYTY	713
Db	195	AIMSDTTNIQRDSQVGTLSYMSPEAF--MCNESDASGNIKCGRPSDIWSLGCILYQMVY	252
Qy	714	GKTPFQIINQISKLHAIIDPNHEIEFPDIPEKDLQDVLCCLKRPDPKQRIPIPELLAHP	773
Db	253	GRTPFADYKTFWAKFKVITDPNHEIMYAPVSNPWLLDLMRCLAWDRNERWRIPQLLQHP	312
Qy	774	YVQIQTHPVNQMAKGTTEEMKYVLGQLVG---LNSPNS	808
Db	313	FLV---PPVPCHSSFYQEQTFKLLQLLISEACINDPES	346

Search completed: September 29, 2004, 16:30:17
Job time : 148 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 16:03:10 ; Search time 27 Seconds
(without alignments)
1621.889 Million cell updates/sec

Title: US-10-081-119-14
Perfect score: 4376
Sequence: 1 MNKVRDIKNKFNEDLTDEL.....GGESHNSSSKTFEKKRGKK 841

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4376	100.0	841	1	TTK HUMAN
2	3165.5	72.3	856	1	TTK_MOUSE
3	673.5	15.4	678	1	MPH1_SCHPO
4	586.5	13.4	764	1	MPS1_YEAST
5	347.5	7.9	772	1	TLK2_MOUSE
6	335	7.7	718	1	TLK2_MOUSE
7	332	7.6	766	1	TLK1_HUMAN
8	328	7.5	766	1	TKL1_MOUSE
9	327	7.5	460	1	CDS1_SCHPO
10	324.5	7.4	576	1	SNF1_SCHPO
11	318.5	7.3	1607	1	M3K4_HUMAN
12	309	7.1	1231	1	YKT3_CAEEL
13	304.5	7.0	1401	1	WIS4_SCHPO
14	303.5	6.9	1597	1	M3K4_MOUSE
15	299	6.8	1478	1	BCK1_YEAST
16	295	6.7	1080	1	NRK1_YEAST
17	294.5	6.7	515	1	STE7_YEAST
18	294.5	6.7	631	1	CNK_MOUSE
19	294.5	6.7	1081	1	KOKO_YEAST
20	294.5	6.7	1142	1	PAK1_YEAST
21	292	6.7	1914	1	KMLS_HUMAN
22	289.5	6.6	544	1	PAK3_MOUSE
23	288.5	6.6	610	1	CDP1_ARATH
24	288.5	6.6	620	1	SNF1_CANAL
25	287.5	6.6	730	1	CDK9_CAEEL
26	287	6.6	988	1	PKK1_SCHPO
27	285.5	6.5	545	1	PAK1_MOUSE
28	285.5	6.5	1906	1	KMLS_CHICK
29	285	6.5	615	1	CNK_RAT
30	282.5	6.5	443	1	ST24_HUMAN
31	282.5	6.5	544	1	PAK3_HUMAN
32	282.5	6.5	646	1	CNK_HUMAN
33	281.5	6.4	491	1	STK3_HUMAN

34	280.5	6.4	544	1	PAK3_RAT	Q62829	rattus norv
35	280	6.4	891	1	KIN1_SCHPO	P22987	schizosacch
36	279.5	6.4	544	1	PAK1_RAT	P35465	rattus norv
37	279	6.4	387	1	KPBG_MOUSE	P07934	mus musculu
38	278	6.4	387	1	KPBG_RAT	P13286	rattus norv
39	278	6.4	586	1	KMS_YEAST	Q03533	saccharomyc
40	277.5	6.3	893	1	KMSR_YEAST	P53739	saccharomyc
41	277	6.3	313	1	NEK6_HUMAN	Q9hc98	homo sapien
42	277	6.3	313	1	NEK6_MOUSE	Q9es70	mus musculu
43	277	6.3	408	1	ST6L_XENLA	Q91819	xenopus lae
44	276.5	6.3	384	1	ARK1_SCHPO	O59790	schizosacch
45	276.5	6.3	511	1	NEK3_MOUSE	Q9r0a5	mus musculu

ALIGNMENTS

RESULT 1

TTK_HUMAN	STANDARD;	PRT;	841 AA.
ID TTK_HUMAN			
AC P33981;			
DT 01-FEB-1994 (Rel. 28, Created)			
DT 01-FEB-1994 (Rel. 28, Last sequence update)			
DT 16-OCT-2001 (Rel. 40, Last annotation update)			
DE Dual specificity protein kinase TTK (EC 2.7.1.1.-) (PYT).			
GN TTK OR MPS1L1.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=92348472; PubMed=1639825;			
RA Mills G.B., Schmandt R., McGill M., Amendola A., Hill M.,			
RA Jacobs K., May C., Rodricks A.-M., Campbell S., Hogg D.;			
RT "Expression of TTK, a novel human protein kinase, is associated with			
RT cell proliferation.";			
RL J. Biol. Chem. 267:16000-16006(1992).			
RN [2]			
RP SEQUENCE OF 509-776 FROM N.A.			
RX MEDLINE=92065863; PubMed=1956325;			
RA Lindberg R.A., Hunter T.;			
RL Unpublished results, cited by:			
RL Hanks S.K., Quinn A.M.;			
RL Meth. Enzymol. 200:38-62(1991).			
CC -I- FUNCTION: PHOSPHORYLATES PROTEINS ON SERINE, THREONINE, AND			
CC TYROSINE. PROBABLY ASSOCIATED WITH CELL PROLIFERATION.			
CC -I- TISSUE SPECIFICITY: PRESENT IN RAPIDLY PROLIFERATING CELL LINES.			
CC -I- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions, as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC -----			
CC EMBL; M86699; AAA61239.1; -.			
CC PIR; A42861; A42861.			
CC HSSP; Q00534; 1BI8.			
CC Genew; HGNC:12401; TTK.			
CC NIM; 604092; -.			
CC GO; GO:0005819; C:spindle; TAS.			
CC GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.			
CC GO; GO:0004713; F:protein-tyrosine kinase activity; TAS.			
CC GO; GO:0007052; P:mitotic spindle assembly; TAS.			
CC GO; GO:0007094; P:mitotic spindle checkpoint; TAS.			
CC GO; GO:0008284; P:positive regulation of cell proliferation; TAS.			
CC GO; GO:0000074; P:regulation of cell cycle; TAS.			
CC InterPro; IPR000719; Prot_kinase.			
CC InterPro; IPR008271; Ser_thr_kinase.			
CC InterPro; IPR002290; Ser_thr_kinase.			

DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Tyrosine-protein kinase.
FT DOMAIN 509 775 PROTEIN_KINASE.
FT NP_BIND 515 523 ATP (BY SIMILARITY).
FT BINDING 537 537 ATP (BY SIMILARITY).
FT ACT_SITE 631 631 BY SIMILARITY.
FT CONFLICT 752 752 L -> V (IN REF. 2).
SQ SEQUENCE 841 AA; 95279 MW; ACDEC2ACD15EF08C CRC64;

Query Match 100.0%; Score 4376; DB 1; Length 841;
Best Local Similarity 100.0%; Pred. No. 1.4e-213;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKVRDIKNKPKNEDLTDELSLNKISADTTDNGTGVNQIMMANNPEDWLSLLKLEKNS 60
DB 1 MNKVRDIKNKPKNEDLTDELSLNKISADTTDNGTGVNQIMMANNPEDWLSLLKLEKNS 60

QY 61 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRPAELKAIQEPDARDYFQMAR 120
DB 61 VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRPAELKAIQEPDARDYFQMAR 120

QY 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALRNLNQKKQLLS 180
DB 121 ANCKKFAFVHISFAQFELSQGNVKKSKQLLQKAVERGAVPLEMLEIALRNLNQKKQLLS 180

QY 181 EEEKNLSASTVLTAEQSFSGSLGHLQNRNNSCDRSGQTKARFLYGENMPPQDAEIGYR 240
DB 181 EEEKNLSASTVLTAEQSFSGSLGHLQNRNNSCDRSGQTKARFLYGENMPPQDAEIGYR 240

QY 241 NSLRQTKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECDLWVPGSKP 300
DB 241 NSLRQTKTKQSCPFGRVPVNLNNSPDCDVKTDDSVVPCFMKQTSRSECDLWVPGSKP 300

QY 301 SGNDSCELNLSQVNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360
DB 301 SGNDSCELNLSQVNSHFKEPLVSDKSSSELIITDSITLKNKTESLLAKLEETKEYQE 360

QY 361 PEVPESNQKQWQAKRKSECINQNPAASSNHQIPELARKVNTQKHHTTEQPVFVSQSKS 420
DB 361 PEVPESNQKQWQAKRKSECINQNPAASSNHQIPELARKVNTQKHHTTEQPVFVSQSKS 420

QY 421 PPISTSKWDFPKSICKTPSSNTLDDYMSCFRTPVVKNDPPACQLSTPYQPCAFQQQQH 480
DB 421 PPISTSKWDFPKSICKTPSSNTLDDYMSCFRTPVVKNDPPACQLSTPYQPCAFQQQQH 480

QY 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSSKVFQVLNEKKQIYAIFYN 540
DB 481 QILATPLQNLQVLASSANECISVKGRIYSILKQIGSGGSSKVFQVLNEKKQIYAIFYN 540

QY 541 LEEADNQTLDSYRNEIAYLNKLQOHSDKIIRLYDYBITDQIYIMVMECNGIDLNSWLKX 600
DB 541 LEEADNQTLDSYRNEIAYLNKLQOHSDKIIRLYDYBITDQIYIMVMECNGIDLNSWLKX 600

QY 601 KSIDPWERKSYWKMLAEAVTHIQHGVHSDKLPANFLIVDGMKLIDFGIANQMPDIT 660
DB 601 KSIDPWERKSYWKMLAEAVTHIQHGVHSDKLPANFLIVDGMKLIDFGIANQMPDIT 660

QY 661 SVVKDSQVGTVMYMPPEAIKDMSSSRRENGKSKSPKSDVWSLGCILYMTYKTPFQQ 720
DB 661 SVVKDSQVGTVMYMPPEAIKDMSSSRRENGKSKSPKSDVWSLGCILYMTYKTPFQQ 720

QY 721 IINQISKLHAIDPNHEIEFPDIPKDLQDLVKCLKRDPKQRISEPELLAHPYVQIQTH 780
DB 721 IINQISKLHAIDPNHEIEFPDIPKDLQDLVKCLKRDPKQRISEPELLAHPYVQIQTH 780

QY 781 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLHYHSGESHSSSKTFEKKRGK 840

DB 781 PVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKTLHYHSGESHSSSKTFEKKRGK 840
QY 841 K 841
DB 841 K 841

RESULT 2
TTK MOUSE
ID TTK MOUSE STANDARD; PRT; 856 AA.
AC P35761;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein kinase TTK (EC 2.7.1.-) (ESK) (PYT).
GN TTK OR ESK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=92269841; PubMed=1375325;
RA Douville E.M.J., Afar D.E.H., Howell B.W., Letwin K., Tannock L.,
RA Ben-David Y., Pawson T., Bell J.C.;
RT "Multiple cDNAs encoding the esk kinase predict transmembrane and
RT intracellular enzyme isoforms";
RL Mol. Cell. Biol. 12:2681-2689(1992).
CC -!- FUNCTION: PHOSPHORYLATES PROTEINS ON SERINE, THREONINE, AND
CC TYROSINE. MAY PLAY SOME ROLE IN THE CONTROL OF CELL PROLIFERATION
CC OR DIFFERENTIATION AND COULD BE INVOLVED IN MODULATING DIFFERENT
CC LEVELS OF SIGNAL TRANSDUCTION PATHWAYS.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P35761-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P35761-2; Sequence=VSP_004875;
CC -!- TISSUE SPECIFICITY: PRESENT IN RAPIDLY PROLIFERATING CELL LINES;
CC HIGH LEVELS IN TESTIS, BONE MARROW, SPLEEN AND THYMUS. LOW LEVELS
CC IN BRAIN, HEART, LUNG AND KIDNEY.
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M86377; AAA37578.1; -.
DR PIR; A44439; A44439.
DR PIR; B44439; B44439.
DR HSP; Q00534; 1BI8.
DR MGD; MGI:1194921; Ttk.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Tyrosine-protein kinase; Alternative splicing; Phosphorylation.
FT DOMAIN 524 790 PROTEIN_KINASE.
FT NP_BIND 530 538 ATP (BY SIMILARITY).
FT BINDING 552 552 ATP (BY SIMILARITY).

FT	ACT SITE	646	646	BY SIMILARITY.
FT	VARSP LIC	296	321	Missing (in isoform 2).
FT				/FTid=Vsp 004875.
SQ	SEQUENCE	856 AA;	96210 MW;	31F8C16C195E6E86 CRC64;
Query Match				
Best Local Similarity 72.3%; Score 3165.5; DB 1; Length 856;				
Matches 622; Conservative 86; Mismatches 107; Indels 51; Gaps 8;				
Qy	1	MNKVRDIKNKFNEDLTDELSLNKISADTTDSGTYNQIMMANNPEDWLSLLKLEKNS	60	
Db	17	MSKVRDIKNKI-NEDCTDELSLKICADHTE--TVNQIMRVGNTPENWLNFLKLEKNS	72	
Qy	61	VPLSDALLNKLIGRYSQAIEALPPDKYQNESFARIQVRFABELKAIQEPDDARDYFQMAR	120	
Db	73	SPLNDDLNKLIGRYSQAIEVLPPDKYQNESFARIQVRLAELKAIQEPDDARDYFQMAR	132	
Qy	121	ANCKKFAFVHISFAQFELSQGNVKKSKOLLQKAVERGAVPLEMLEIALRNLNLOKKQLLS	180	
Db	133	ENCKKFAFVHVSFAQFELSQGNLKKSEQLLHKAVETGAVPLQMLETAMRNHLQKKQLLP	192	
Qy	181	EECKNLSASTVLTAESEFSGLHLQNRNNSCDRSGQTTKARFLYGENMPPQDAEIGYR	240	
Db	193	EEDKKSVSASTVLSAQEPFSSSLGNVQNRISCESRGQAGAARVLYGENLPPQDAEVRHQ	252	
Qy	241	NSLRQTNKTKQSCPGFGRVPVNLNNSPCDCKVTDDSV-----VPCFMKR	283	
Db	253	NPFKQTHAAKRCPCFGRVPVNLNNSPDFVYKTDSSAVTQLTTRLALSSVPLPYVTCLLHL	312	
Qy	284	Q-----TSRSECRDLVPGSKPSGNDSCELRNLSKVQNSHFKEPLVSDEKSELIIIT	335	
Db	313	QLLALAGLAGSGGPRDAILPGSRPRGSDSYELRGLKPIQTIYLYKDSLVSNKSEL-MS	371	
Qy	336	DSITLKNKTESSLLAKLEETKEYQEPEVPESNQKQWAKRKBECINQNPAASSNHWQIPE	395	
Db	372	DLIALKSKTDSS-LTKLEETK----PEIARRPMQWQSTRKEPCVFNQPAAFAPLHVPD	426	
Qy	396	LARKVNTQKHTTFEQPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCRTPVV	455	
Db	427	VTPKAD-----KESPPISVPKWLDPKSA CETPSSSSLLDDYMKCKTPVV	470	
Qy	456	KNDFPACQLSTPYGQACFQQQHQIILATPLQNLQVLASSANECISVKGRYISILKQI	515	
Db	471	KNDFPACPSSTPYSQLARLQQQQGLSTPLQSLQISGSSSINECISVNGRIYSILKQI	530	
Qy	516	GSGSSKVFQVLENEKQIYAIKYVNLLEEADNQTLDSYRNEIAYLNKLQOHSKIRLYDY	575	
Db	531	GSGSSKVFQVLENEKQINA KYVNLLEDADNSQTIESYRNEIAPLNKLQOHSKIRLYDY	590	
Qy	576	EITDQYIYMMECGNIDLSWLKCKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPA	635	
Db	591	EITEQYIYMMECGNIDLSWLKCKSINPWERKSYWKNMLEAVHIIHQHGIVHSDLKPA	650	
Qy	636	NFLIVDGMKLIDFGIANQMOPDTTSVVKDSQVGTVNYMPPEAIKDMSSSRENGSKSKI	695	
Db	651	NEVIVDGMKLIDFGIANQMOPDTTSIVKDSQVGTVNYMAPEAIRDMSSSRENSKIRTKV	710	
Qy	696	SPKSDVWSLGCILYMYTYGKTPFQIINQISKLHAIIDPNHEIBPDPDIPEKDLQDVLKCC	755	
Db	711	SPRSDVWSLGCILYMYTYGRTPFQHIINQVSKLHAIINPAHEIBPPEISEKDLRDLKCC	770	
Qy	756	LKRDPKQKISIPPELLAHPYVQIQTHFPVNQMAKGTTEEMKYVLGQLVGLNSPNSILKAAKT	815	
Db	771	LVRNPKERISIPELLTHPYVQIQPHPGSQMARGATDEMKYVLGQLVGLNSPNSILKTAKT	830	
Qy	816	LYEHYSGGESHNSSSSKTFEKKRGKK	841	
Db	831	LYERYNCGEQDSSSSKTFDKKRERK	856	

RESULT 3
MPH1_SCHPO
ID MPH1_SCHPO STANDARD; PRT; 678 AA.

O94235; Q9P7Z5; Q9URW0;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Serine/threonine protein kinase mph1 (EC 2.7.1.37).
GN MPH1 OR SPBC106.01 OR SPBC243.01 OR SPBC1271.16C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=98264860; PubMed=9601094;
He X., Jones M.H., Winey M., Sazer S.;
RT "Mph1, a member of the Mps1-like family of dual specificity protein
RT kinases, is required for the spindle checkpoint in *S. pombe*.";
RL J. Cell Sci. 111:1635-1647(1998).
[2]
SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21948401; PubMed=11859360;
Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Involved in the regulation of the onset of mitosis.
CC Involved in a pathway that coordinates cell proliferation and
CC differentiation.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AF020705; AAD01648.1; -
DR EMBL; AL133362; CAB62431.1; -
DR EMBL; AL034353; CAB72266.1; -
DR EMBL; AL110295; CAB53717.1; -
DR PIR; T43539; T43539.
DR GeneDB_Spombe; SPBC106.01; -
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR GO; GO:0000074; P:regulation of cell cycle; ISS.
DR GO; GO:0007165; P:signal transduction; ISS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR

FT	ACT SITE	563	563	BY SIMILARITY.	
FT	MUTAGEN	580	580	D->A: LOSS OF ACTIVITY.	
FT	CONFLICT	146	146	A -> S (IN REF. 1).	
FT	CONFLICT	211	213	RRE -> TKR (IN REF. 1).	
SQ	SEQUENCE	764 AA;	86827 MW;	26B171DFF8B8EB4D CRC64;	
Query Match 13.4%; Score 586.5; DB 1; Length 764;					
Best Local Similarity 26.1%; Pred. No. 1.3e-22;					
Matches 188; Conservative 140; Mismatches 230; Indels 161; Gaps 27;					
QY	178	LLSEEEKNLSASTVLTAESES	-----GSLGHLQNRNNSCDSRGQT	219	
Db	40	LLSHTK--TSASEILSSHNNDKIANRLPEEMDRSSSRSHPPSPMGNLTSGHTSTSSH-ST	96		
QY	220	TKARFLYGEN------MPPQDAEIGYRNS	-----LRQ-----	245	
Db	97	LFGRYLRNNHQTSMTMTNTSDIEINVGNGLDKSFERIRNLRQNMKEDITAKYAERRSKRF	156		
QY	246	--TNKTKQSCPFGRVPV-----NLLNSPDCDVKTDDSV--VPCFMKRQTSRSECRDLVV	295		
Db	157	LISNRTTKLGPAKRAMTLTNIFDEVPNSNPQIPINARETVELPLEDSHQTNFKERRE--	213		
QY	296	PGSKPSGNDSCELNLSVQNSHFKE-----PLVSD-----	EKSS 330		
Db	214	-----NTDYDSIDFGDLNPIQ--YIKKHNLPSTDPLISQIYFDKQREENQAALRKHSSR	267		
QY	331	ELIITDSITLKNKTESLLAKLEETKEYQPEVPESNQKQWQAKRKSECINQNPAASSNH	390		
Db	268	ELLYKSRSSSSLSNNLLAN-----KDNSITSNNGS--QPRK--VSTGSSSSKSS	315		
QY	391	WQIPELARKVNTQKHTTPEQPVFSVSKQPPISTSKWFD--PKSICKTPSSN-----	441		
Db	316	IEIRRALKENIDTSNNSNFNSPIHKIYK---GISRNKDSSEKREVLNRNISINAHADNL	372		
QY	442	-----TLDDVMSCFRTPVVKNDFFPACQLSTPYGQPAFCQQQHQI-LATFLQNL	490		
Db	373	LQQENKRLKRLDD-----AITNENINSKNLEVFYHRPAPKPPVTKKVEIPEAKSA	424		
QY	491	QVLASSANECISVKGRYISILKQIGSGGSKVQVFLNEKKQIYAIKYVNLLEEADNQTLD	550		
Db	425	SL---SNNRNIITVNDQYKEIKELLGRGSSRVYKVGSGNRVYALKRVSFDAFDDSSID	481		
QY	551	SYRNEIAYNLKQHSKIRLYDYEITDQYIMVMECGNIDLSWLKKKS--IDPWER	608		
Db	482	GFGKEIELEKLDQK-RVIQLLDYEMGDGLLYLIMECGDHDLSQILNQRSGMPLDFNFV	540		
QY	609	KSYWKNMLEAVHTIHQGHVHSDLKPANFLIVDGMKLIDFGIANQMPDPTTSVVKDSQV	668		
Db	541	RFYTKEMLLCIKVVDAGIVHSDLKPANFVLVKILKIIDFGIANAVPEHTVNIYRETOI	600		
QY	669	GTVMYMPPEAIKDM-----SSSRENGSKSKISPKSDWSLGLCYLYMTYKTPFQIINQ	724		
Db	601	GTPNYMAPEALVAMNYTQNSENQHEGNKWKVGRPSDMWSCGCIYIYQMIYKPPYGSFQOQ	660		
QY	725	ISKLHAIIDPNHEIEFPD-----IPEKDLQDVLKCLKRDPKQRISIPELLAHPYVQ	776		
Db	661	-NRLLAIMPDPVKIPFPEHTSNNEKIP-KSAIELMKACLYRNPDKRWTVDKVLSSTFLQ	717		
RESULT 5					
TLK2	HUMAN	STANDARD;	PRT;	772 AA.	
ID	TLK2	HUMAN	Q86UE8; Q9UKI7; Q9Y4F7;		
AC	Q86UE8; Q9UKI7; Q9Y4F7;				
DT	10-OCT-2003 (Rel. 42, Created)				
DT	10-OCT-2003 (Rel. 42, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Serine/threonine-protein kinase tousled-like 2 (EC 2.7.1.37) (Tousled-				
DE	like kinase 2) (PKU-alpha).				
GN	TLK2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				

RN	[1]	SEQUENCE FROM N.A. (ISOFORM 3), FUNCTION, SUBCELLULAR LOCATION, AND
RP		TISSUE SPECIFICITY.
RC		TISSUE=Placenta, and Testis;
RX		MEDLINE=98087437; PubMed=9427565;
RA		Yamakawa A., Kameoka Y., Hashimoto K., Yoshitake Y., Nishikawa K.,
RA		Tanihara K., Date T.;
RT		"cDNA cloning and chromosomal mapping of genes encoding novel protein
RT		kinases termed PKU-alpha and PKU-beta, which have nuclear localization
RT		signal.";
RL		Gene 202:193-201(1997).
RN	[2]	SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, MUTAGENESIS OF ASP-592,
RP		SUBCELLULAR LOCATION, INTERACTION WITH TLK1, AND ENZYME REGULATION.
RC		TISSUE=Placenta;
RX		MEDLINE=99452775; PubMed=10523312;
RA		Sillje H.H.W., Takahashi K., Tanaka K., Van Houwe G., Nigg E.A.;
RT		"Mammalian homologues of the plant tousled gene code for cell-cycle-
RT		regulated kinases with maximal activities linked to ongoing DNA
RT		replication.";
RL		EMBO J. 18:5691-5702(1999).
RN	[3]	SEQUENCE FROM N.A. (ISOFORM 1).
RP		TISSUE=Testis;
RX		MEDLINE=22388257; PubMed=12477932;
RA		Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA		Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA		Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA		Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA		Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA		Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA		Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA		Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA		Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA		Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA		Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA		Fahey J., Helton E., Kettenan M., Madan A., Rodrigues S., Sanchez A.,
RA		Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA		Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA		Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA		Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA		Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT		"Generation and initial analysis of more than 15,000 full-length human
RT		and mouse cDNA sequences.";
RL		Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[4]	FUNCTION, AND ENZYME REGULATION.
RP		MEDLINE=22546651; PubMed=12660173;
RX		Groth A., Lukas J., Nigg E.A., Sillje H.H.W., Wernstedt C., Bartek J.,
RA		Hansen K.;
RA		"Human Tousled like kinases are targeted by an ATM- and Chk1-dependent
RT		DNA damage checkpoint.";
RT		EMBO J. 22:1676-1687(2003).
CC	-!	FUNCTION: Rapidly and transiently inhibited by phosphorylation
CC		following the generation of DNA double-stranded breaks during S-
CC		phase. This is cell cycle checkpoint and ATM-pathway dependent and
CC		appears to regulate processes involved in chromatin
CC		assembly.
CC	-!	CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC	-!	COFACTOR: Magnesium.
CC	-!	ENZYME REGULATION: Cell-cycle regulated, maximal activity in S-
CC		phase. Inactivated by phosphorylation at Ser-749, potentially by
CC		CHK1.
CC	-!	SUBUNIT: Heterodimerizes with TLK1.
CC	-!	SUBCELLULAR LOCATION: Nuclear.
CC	-!	ALTERNATIVE PRODUCTS:
CC		Event=Alternative splicing; Named isoforms=3;
CC		Name=1;
CC		ISOId=Q86UE8-1; Sequence=Displayed;
CC		Note=No experimental confirmation available;
CC		Name=2;
CC		ISOId=Q86UE8-2; Sequence=VSP_050573;
CC		Name=3;

QY 321 EPLVDSKSELITDITLKNKTESSLLAKLEETKEYQPEVPESNQKQWQAKRKSECI 380
Db 278 EQW-----TDGYAFQN-----LIKQERINSOREI--ERQKMLAKRPPAM 318
QY 381 NQNPAASSNHQIPELARKVNTQKHTTEQPVFSVSKQSPPISTSKWFDPKSICKTPSS 440
Db 319 QGAPPATNEQKQ-----RKSNTN-----GAENETLTLAEYHEQEIEFKRLG 360
QY 441 NTLDDYMSCFRTPVVKNDFFPACQLSTPYGQPACFQQQHQILATPLQNLQVLASSANE 500
Db 361 H-----LKKEEAIIQA-ELERLERVRNLHIRE 386
QY 501 CISVKGR-----IYSILKQIGSGGSSKVFQVLNKKQIY-AIKVYNLEEA-D 545
Db 387 LKRIHNEDNSQFQDHPPTLNDRYLLHLLGRGGFSEVYKAFDLTEQRYVAVKIHQLNKNWR 446
QY 546 NQTLDSYRNEIAYLNKLQOHS--KIIRLYDYEI--TDQYIYMMECGNIDLSWLKXK 601
Db 447 DEKENYHKHACREYRIHKELDPRIKLYDYFSLDTSFCTVLEYCEGNDLDFYLKQHK 506
QY 602 SIDPWERKSYKMNMLEAVHTIHO--HGIIVHSDLKPNFLIVDGM----LKLIDFGIANQM 655
Db 507 LMSEKEARSIIQIVNALKYLNKPIIHYDLKPGNILLVNGTACGEIKITDGLSKIM 566
QY 656 QPDTTSVVKDSQV-----GTVYMPPEAIKDMSSRENGKSKISPKSDVMSLGCILYY 710
Db 567 DDDSYNSVDGMELTSGAGTYWLPPECF-----VVGKEPPKISNKVDVMSVGVIFYQ 619
QY 711 MTYGTPTF-----QQIINQISKLHAIIDPNHEIEFPDIP--EKDLQDLVKCKLRDPKQ 762
Db 620 CLYGRKPFQHNQSQDDILOENTILKAT-----EVQFPKPVVTPPEAKAFIRCLAYRKED 674
QY 763 RISIPELLAHYPVQIQTHPVNQMAKGTTEEMKYVLGVLNLSNPSILKAAKTLYEHSYG 822
Db 675 RIDVQQLACDPY--LLPH-----IRKSIVSTSSPAGAAIAS-----TS 709
QY 823 GESHNSSSS 831
Db 710 GASNNSSSN 718

RESULT 7

TLK1_HUMAN
ID TLK1_HUMAN STANDARD; PRT; 766 AA.
AC Q9UKI8; Q14150; Q8N591; Q9NYH2; Q9Y4F6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase tousled-like 1 (EC 2.7.1.37) (Tousled-like kinase 1) (PKU-beta).
GN TLK1 OR KIAA0137.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Placenta, and Testis;
RX MEDLINE=98087437; PubMed=9427565;
RA Yamakawa A., Kameoka Y., Hashimoto K., Yoshitake Y., Nishikawa K., Tanihara K., Date T.;
RT "cDNA cloning and chromosomal mapping of genes encoding novel protein kinases termed PKU-alpha and PKU-beta, which have nuclear localization signal.";
RL Gene 202:193-201(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, MUTAGENESIS OF ASP-586, SUBCELLULAR LOCATION, INTERACTION WITH TLK2, AND ENZYME REGULATION.
RC TISSUE=Placenta;
RX MEDLINE=99452775; PubMed=10523312;
RA Sillje H.H.W., Takahashi K., Tanaka K., Van Houwe G., Nigg E.A.;
RT "Mammalian homologues of the plant tousled gene code for cell-cycle-

regulated kinases with maximal activities linked to ongoing DNA replication.";
RL EMBO J. 18:5691-5702(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 3), FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Placenta;
RX MEDLINE=20056088; PubMed=10588641;
RA Cabanils J.-P., Ravichandran V., Roche P.A.;
RT "Phosphorylation of SNAP-23 by the novel kinase SNAK regulates t-SNARE complex assembly.";
RL Mol. Biol. Cell 10:4033-4041(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Bone marrow;
RX MEDLINE=96127530; PubMed=8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:167-174(1995).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP FUNCTION, MUTAGENESIS OF SER-743, PHOSPHORYLATION OF SER-743, AND ENZYME REGULATION.
RX MEDLINE=22546651; PubMed=12660173;
RA Groth A., Lukas J., Nigg E.A., Sillje H.H.W., Wernstedt C., Bartek J., Hansen K.;
RT "Human Tousled like kinases are targeted by an ATM- and Chk1-dependent DNA damage checkpoint.";
RL EMBO J. 22:1676-1687(2003).
CC -!- FUNCTION: Rapidly and transiently inhibited by phosphorylation following the generation of DNA double-stranded breaks during S-phase. This is cell cycle checkpoint and ATM-pathway dependent and appears to regulate processes involved in chromatin assembly. Isoform 3 phosphorylates and enhances the stability of the t-SNARE SNAP23, augmenting its assembly with syntaxin.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- COFACTOR: Magnesium.
CC -!- ENZYME REGULATION: Cell-cycle regulated, maximal activity in S-phase. Inactivated by phosphorylation at Ser-743, potentially by CHK1.
CC -!- SUBUNIT: Heterodimerizes with TLK2.
CC -!- SUBCELLULAR LOCATION: Nuclear. Isoform 3 appears to be cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9UKI8-1; Sequence=Displayed;
CC Name=2;

CC IsoId=Q9UKI8-2; Sequence=VSP_050571;
CC Name=3; Synonyms=SNAK;
CC IsoId=Q9UKI8-3; Sequence=VSP_050570;
CC TISSUE SPECIFICITY: Widely expressed. Present in fetal placenta,
CC liver, kidney and pancreas but not heart or skeletal muscle. Also
CC found in adult cell lines. Isoform 3 is ubiquitously expressed in
CC all tissues examined.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB004885; BAA20562.1; -;
DR EMBL; AF162666; AAF03094.1; ALT_INIT.
DR EMBL; AF246219; AAF71263.1; -;
DR EMBL; D50927; BAA09486.2; ALT_INIT.
DR EMBL; BC032657; AAH32657.1; -;
DR Genew; HGNC:11841; TLK1.
DR GO; GO:0005634; C:nucleus; IEP.
DR GO; GO:0005524; F:ATP binding; IDA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
DR GO; GO:006886; P:intracellular protein transport; IDA.
DR GO; GO:0007242; P:intracellular signaling cascade; IDA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0001672; P:regulation of chromatin assembly/disassembly; IDA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SMO0220; S_TKC_1.
DR SMART; SMO0219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Chromatin regulator; ATP-binding; Serine/threonine-protein kinase;
KW transferase; Cell cycle; DNA damage; Nuclear protein; Coiled coil;
KW phosphorylation; Alternative splicing.
FT DOMAIN 230 281 COILED COIL (POTENTIAL).
FT DOMAIN 397 445 COILED COIL (POTENTIAL).
FT DOMAIN 456 734 PROTEIN KINASE.
FT NP_BIND 462 470 ATP (BY SIMILARITY).
FT BINDING 485 485 ATP (BY SIMILARITY).
FT ACT_SITE 586 586 PHOSPHORYLATION.
FT MOD_RES 743 743 Missing (in isoform 3).
FT VARSPLIC 1 217 /FTId=VSP_050570.
FT VARSPLIC 136 136 G -> GPNLPVFQSLAYWEMGRIAGG (in isoform
FT 2).
FT /FTId=VSP_050571.
FT D->A: LOSS OF KINASE ACTIVITY.
FT S->A: LOSS OF KINASE INHIBITION IN
FT RESPONSE TO DNA DAMAGE.
FT S->E: LOSS OF KINASE INHIBITION IN
FT RESPONSE TO DNA DAMAGE.
FT S->D: LOSS OF KINASE INHIBITION IN
FT RESPONSE TO DNA DAMAGE.
FT S -> T (IN REF. 2).
FT G -> E (IN REF. 1).
FT Q -> L (IN REF. 1).
FT E -> D (IN REF. 1).
FT E -> G (IN REF. 1).
FT N -> H (IN REF. 1).
FT Y -> D (IN REF. 1).
FT Y -> S (IN REF. 1).
FT D -> V (IN REF. 1).
FT F -> Y (IN REF. 1).
FT CONFLICT 88 88
FT CONFLICT 102 102
FT CONFLICT 230 230
FT CONFLICT 261 261
FT CONFLICT 416 416
FT CONFLICT 439 439
FT CONFLICT 471 471
FT CONFLICT 477 477
FT CONFLICT 625 625
FT CONFLICT 665 665

FT CONFLICT 730 730 N -> C (IN REF. 1).
SQ SEQUENCE 766 AA; 86699 MW; A65AE0A2A7C7FF2F CRC64;

Query Match 7.6%; Score 332; DB 1; Length 766;
Best Local Similarity 21.3%; Pred. No. 8.9e-10;
Matches 171; Conservative 124; Mismatches 268; Indels 238; Gaps 34;

QY 132 SFAFELS--QGNVKSQQLQKAVERGAFLVEMLEIALRNLNLCQKQLLSEEEKNL 189
Db 15 SWSQLSTPTPGSAAARSLLNHTPPSGR-PREGAMDELHSLDPRRQELLEARFTGVASG 73

QY 190 STVLT-----AQESFGSLGHLQNRNN-----SCDSRGQTTKARFLYG 227
Db 74 STGSTGSCSVGAKASTNNSSNHSF-GSLGSLSDKESETPEKKQSESSRGRKRAE---N 129

QY 228 ENMPQDAEIGYRNSLRQTNKTKQSCPFGRVFNLLNSPDCDVKTDDSVVPCFMKQTSR 287
Db 130 QNESSQKSGIGRGH-----KISD-----YFEYQ--- 153

QY 288 SECDLVVPGSKPSGNDSCELRN---LKSQNSHF-----KEP 322
Db 154 -----GGGSSPVRGIPPAIRSPONSHSHSTPSSSVRPNSPSTALAFGDHP 200

QY 323 LVSDEK-SSELIITDSITLKNKTESLLAKLETK---EYQEPEVP----- 364
Db 201 IVQPKQLSPFKIIQTDLTMLK-----LALESNKIQDLEKKEGRIDLLRANCDLRRQI 253

QY 365 ESNOK--QWQAKRKSECIN-----QNPAASS 388
Db 254 DEQKLEKYEYKERLNKNCISMSKLLIEKSTQEKLSREKSMQDLRLGHFTTVRHGASFT 313

QY 389 NHW-----QIPELARKVN-TEQKHTTFEQPVFSVSKQSPPISTSKWDFPKSICKTPSSNTL 443
Db 314 EQWTDGFAQNLVKQEWVNVQOREDIERQKLLAKRKPPTANNS-----QAPSTNSE 365

QY 444 DDYM-----SCFRTPVVKNDFPPACQLSTPVCQACFQ-----QQHQILATPLQNLQ 491
Db 366 PKQRKNKAVNGAENDPFVRPNLPQLLTLAEYHQBEEIFKLRLGHLKKEEAEIQAELE 425

QY 492 VLAASSANECISVK-----GRIYSILKQIGSGGSKVQVFLNEKKQIY-AIK 537
Db 426 RVRNLHIRELKRINNEDNSQFKDHTLNERYLLHLLGRGGFSEVYKAFDLYEQRYAAVK 485

QY 538 YVNLEEA-DNQTLDYRNEIAYLNKLOQHS--KIIRLYDYEI--TDQYIYVMECEGND 592
Db 486 IHQLNKSWRDEKKNYKHACREYRIHKELDHPRIVKLYDYFSLDPTDFTCTVLEYCEGND 545

QY 593 LNSWLKKKSIDPWERKSYWKNMLEAVHTIQ--HGIVHSDLKPANFLIVDGM---LKL 646
Db 546 LDFYLKQHKIMSEKEARSIVMIVNALRYLNEIKPPIIHYDLKPGNILLVDGTACGEIKI 605

QY 647 IDFGIANQMQPDTTSV---VKDSQVGTVNVMPPEAIKDMSSRENGSKSKISPKSDVW 702
Db 606 TDFGLSKIMDDDSYGVGDMDLTSQAGAGTYWYLPPECF-----VVGKEPPKISNKVDVW 658

QY 703 SLGCILYMYTYGKTPF-----QQIINQISKLHAIIDPNHEIEFPDIP--EKDLQDVLKC 754
Db 659 SVGVIFFFQCLYGRKPFHGNQSQDILQENTILKAT-----EVQFPVKPVVSSEAKAFIR 713

QY 755 CLKRDPKQRISSIPPELLAHPYV 775
Db 714 CLAYKEDRFDVHQLANDPYL 734

RESULT 8
TKL1 MOUSE
ID TKL1 MOUSE STANDARD; PRT; 766 AA.
AC Q8C0V0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase tousled-like 1 (EC 2.7.1.37) (Tousled-
DE like kinase 1).


```
QY 293 LVVPGSKPSGNDSCELRNLSKVQNSHFKEPLVSDKSSSELI----- 334
Db 258 KLEKYKERLN-KCISMSKLLIEKSTQEKLSREKSMQDRRLRGHFTTVRHGASFTQW 316
QY 335 TDSITLKNKTESSLLAKLEETKEYQEPVPSNQKQWAKRSEKSCINQNPAASSNHQIP 394
Db 317 TDGFAQN-----LVKQEWVNOQREDI--ERQKLLGKRKPTANNSQAPATN----- 363
QY 395 ELARKVNTQXHTTTEQPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCRTPV 454
Db 364 -----SEAKQR-----KTKAVNGAEN-----DPF 382
QY 455 VKNDPPACQLSTPYQCPACFQ-----QQHQILATPLQNLQVLASSANECISVK--- 505
Db 383 VRPNLPOLLTLAEVHQEIEFKRLGLHLKKEAEIAQELERLVRNHLHIRELKRINNED 442
QY 506 -----GRIYSILKQIGSGGSKVQVLNEKKQIY-AIKYVNLEEA-DNQTLDSYR 553
Db 443 NSQFKDHPHTLNERYLLHLGRGGFSEVYKAFDLYEQRYAAVKIHLNKSQRDEKKNYH 502
QY 554 NEIAYLNKLQHS--KIIRLYDYEI--TDQYIYMVECGNIDLSWLKKSIDPWERK 609
Db 503 KHACREYRIHKELDHPRIVKLYDYFSLDTFTCTVLEYCEGNDLDFYLKQHKLMSEKEAR 562
QY 610 SYWKNMLEAVHTIQ--HGIVHSDLKPANFLIVDGM---LKLIDFGIANQMOPDTTSV- 662
Db 563 SIVMQIVNALRYLNEIKPPIIHYDLKPGNILLVDGTACGEIKITDFGLSKIMDDSYGVD 622
QY 663 ---VKDSQVGTVNYMPPEAIKDMSSSRENGSKSKISPKSDVWSLGLCILYMYTGKTPF- 718
Db 623 GMDLTSQAGTYWYLPPECF-----VVGKEPPKISNKVDVWSVGVIFFQCLYGRKPF 675
QY 719 -----QIINQISKLHAIIDPNHEIEFPDIP--EKDLQDLVKCLKRDPKQRISIPELLA 771
Db 676 HMQSQDILQENTILKAT-----EVQFPVKPVVSSEAKAFIRCLAYRKEDRFDVHQLAN 730
QY 772 HPYV 775
Db 731 DPYL 734

RESULT 9
CDS1_SCHPO
ID CDS1_SCHPO STANDARD; PRT; 460 AA.
AC Q09170; O42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein kinase cds1 (EC 2.7.1.37) (Checkpoint kinase cds1).
GN CDS1 OR SPCC18B5.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=95240713; PubMed=7723827;
RA Murakami H., Okayama H.;
RT "A kinase from fission yeast responsible for blocking mitosis in S
RT phase.";
RL Nature 374:817-819(1995).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND PHOSPHORYLATION.
RC STRAIN=972;
RX MEDLINE=98119835; PubMed=9450932;
RA Lindsay H.D., Griffiths D.J.F., Edwards R.J., Christensen P.U.,
RA Murray J.M., Osman F., Walworth N., Carr A.M.;
RT "S-phase-specific activation of Cds1 kinase defines a subpathway of
RT the checkpoint response in Schizosaccharomyces pombe.";
RL Genes Dev. 12:382-395(1998).
RN [3]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Mule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsborg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Has a role in the DNA replication-monitoring S/G2
CC checkpoint system. It is responsible for blocking mitosis in the
CC S phase. It monitors DNA synthesis by interacting with DNA
CC polymerase alpha and sends a signal to block the onset of mitosis
CC while DNA synthesis is in progress.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with rad26.
CC -!- PTM: Autophosphorylated.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDS1
CC subfamily.
CC -!- SIMILARITY: Contains 1 FHA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X85040; CAA59410.1; -.
CC EMBL; AJ222869; CAA11019.1; -.
CC EMBL; AL109736; CAB52158.1; -.
CC PIR; S58882; S58882.
CC HSSP; Q00534; 1BI7.
CC GeneDB_SPombe; SPCC18B5.11C; -.
CC InterPro; IPR000253; FHA.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR008984; SMAD FHA.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00498; FHA; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00240; FHA; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50006; FHA_DOMAIN; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Cell cycle; Phosphorylation.
KW DOMAIN 60 116
FT
```



```
FT DOMAIN 167 433 PROTEIN KINASE.
FT NP BIND 173 181 ATP (BY SIMILARITY).
FT BINDING 196 196 ATP (BY SIMILARITY).
FT ACT SITE 294 294 BY SIMILARITY.
FT CONFLICT 61 61 R -> G (IN REF. 1).
FT CONFLICT 238 238 F -> C (IN REF. 1).
SQ SEQUENCE 460 AA; 52014 MW; 4CEB963D3376DB54 CRC64;

Query Match 7.5%; Score 327; DB 1; Length 460;
Best Local Similarity 24.4%; Pred. No. 8.6e-10;
Matches 122; Conservative 95; Mismatches 196; Indels 88; Gaps 22;

QY 322 PL-VSDEKSSSELIITDSITLKNKTESLLAKLEETKEYQEPEVPEGNQKQWAKRKSEC- 379
Db 15 PLHVSQNTAKQVNVNNEVFMK-----LVMTMLDGKTEVIPLETTDVHNGFWFRGRHKSCE 69
QY 380 -INQNPAASNHWQIPELARKVNTQKHTTTEQPVFSVSKQSPPISTSKWFDPKSICKTP 438
Db 70 VVLNGPRVSNFHFIEYQGHNRNDSSENVVF-----LHDHS---SNCTFLNFERLAK-- 118
QY 439 SSNTL-----DDYMSCFRTPVVRKNDPPACQLSTPYQOPACFQQQHQILATPLQNLQVLA 494
Db 119 NSRTILSNGDEIRIGLGP--KDEISFLCQVPVXHSR----- 153
QY 495 SSSANECISVKGRIYSILKQIGSGSSKVFQVLNEKKQI-----YAIKYVN-----LEE 544
Db 154 -DSQKNMKSNSHYEIRITLGS-----TFAVVKLAVEVNSGKYAIKINKRILLTSS 208
QY 545 DNQTLDSYRNEIAYLNKLQHSKDIIRLYDYEITDQIYVMVE-CGNIDLNSWLKSKSI 603
Db 209 EKRATEMFQREIDILKSL--HHPGVVQCHEIFEENDELFIWVEYVEGGDLMDFLIANGSI 266
QY 604 DPWERKSYWKNMLEAVHTIHQHGIVHSDLKPANFLIV-DGMLKLIDFGIANQMOPDTTSV 662
Db 267 DEQDCKPLKQLLETLLHLHKQGVTHRDIKPENILITNDFHLKISDFGLAKVIH--GTGT 324
QY 663 VKDSQVGTVMYMPPEAIKDMSSSRENGKSKSISKPSKDVWSLGLCILIYMTYKTPF---- 718
Db 325 FLETFCGTNGYLAPEVLKSKNVNLDGG-----YDDKVDIWSLGCVLVYMLTASIPFSSS 379
QY 719 -QQIINQISKLHAIIDPNHEIEFPDPIPEKLOQVLCCKLRDPKQRISEIPELLAHP-YVQ 776
Db 380 QAKCIELISKGAYPIEPLLENE---ISEGI-DLNRMLNLEINPEKRISESEALQHPWFYT 435
QY 777 IQTH-----PVNQMAKGTTEEM 793
Db 436 VSTHEHRTPPSSSEHEATEQL 456
```

```
RESULT 10
SNF1_SCHPO STANDARD; PRT; 576 AA.
AC O74536;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SNF1-like protein kinase (EC 2.7.1.-).
GN SPCC74.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
```

```
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC subfamily.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL031543; CAA20833.1; --
CC PIR; T41587; T41587.
CC HSSP; Q63450; 1A06.
CC GeneDB SPombe; SPCC74.03c; --
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR000449; UBA_domain.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50030; UBA; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 34 285 PROTEIN KINASE.
FT DOMAIN 304 345 UBA.
FT NP BIND 40 48 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT SITE 156 156 BY SIMILARITY.
SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;
```

Query Match 7.4%; Score 324.5; DB 1; Length 576;
Best Local Similarity 29.3%; Pred. No. 1.5e-09;
Matches 102; Conservative 67; Mismatches 132; Indels 47; Gaps 15;

```
QY 509 YSILKQIGSGSSKVFQVLNEK-KQIYAIKYVN---LEEADNQTLDYSRNEIAYLNKLQ 564
Db 34 YIIRETLGSGFGKVKLATHTYKQKVALKFISRLKSKDMHM--RVERISYL-KLLR 90
QY 565 HSDKIIRLYDYEITDQIYVMVECGNIDLNSWLKSKSIDPWERKSYWKNMLEAVHTIHQ 624
Db 91 H-PHIILYDVITPTDIDVMVIEYAGGELFDYIVKPKMTDEGRFFQIICAIEYCHR 149
QY 625 HGIVHSDLKPANFLIVDGM-LKLIDFGIANQMOPDTTSVVKDSQVGTVMYMPPEAIKMS 683
Db 150 HKIVHRDLKPENLLDDNLNVKIADFGLSNIM---TDGNFLKTSKSGSPNVAPEVI---- 202
QY 684 SSRENGKSKSISKPSKDVWSLGLCILIYMTYKTPFQQ--IINQISKLHAIIDPNHEIEFP 741
```

Db 203 ----NGKLYA--GPEVDVWSCGIVLYVMLVGRLPFDDEFIPNLFKVNCSV-----YVMP 251

QY 742 DIPKDLQDVLKCCCKPDKQKORISIPELLAHPYVQIQ----THPVNQMAKGTTEEMKYV- 796

Db 252 DFLSPGAQSLRRMIVADPMQRITIQEIRRDWFNVNLPDLRLPMEV-QGSYADSRIVS 310

QY 797 -LGQLVGLNSP-----NSILKAAKTLVEHYSGGESHNSSSSK 832

Db 311 KLGEAMGFSEDIYVEALRSDENNEVKEAYNLLHENQVIOEKSHLSKSK 358

RESULT 11

M3K4_HUMAN

ID M3K4_HUMAN STANDARD; PRT; 1607 AA.

AC Q9Y6R4; Q92612;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Mitogen-activated protein kinase kinase 4 (EC 2.7.1.1--)

DE (MAPK/ERK kinase 4) (MEK kinase 4) (MEKK 4) (MAP three kinase 1).

DE 1).

GN MAP3K4 OR MAPKKK4 OR MEKK4 OR MTK1 OR KIAA0213.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND MUTAGENESIS OF LYS-1371.

RC TISSUE=Fetal liver, and Skeletal muscle;

RX MEDLINE=97449143; PubMed=9305639;

RA Takekawa M., Posas F., Saito H.;

RT "A human homolog of the yeast Ssk2/Ssk22 MAP kinase kinase kinases,

RT MTK1, mediates stress-induced activation of the p38 and JNK

RT pathways.";

RL EMBO J. 16:4973-4982(1997).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM B).

RA Ohara O., Nagase T., Kikuno R., Nomura N.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 68-1607 FROM N.A. (ISOFORM B).

RC TISSUE=Bone marrow;

RX MEDLINE=97191544; PubMed=9039502;

RA Nagase T., Seki N., Ishikawa K.-I., Chira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RT "Prediction of the coding sequences of unidentified human genes. VI.

RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by

RT analysis of cDNA clones from cell line KG-1 and brain.";

RL DNA Res. 3:321-329(1996).

CC -!- FUNCTION: ACTIVATES THE CSBP2, P38 AND JNK MAPK PATHWAYS, BUT NOT

CC THE ERK PATHWAY. SPECIFICALLY PHOSPHORYLATES AND ACTIVATES MAP2K4

CC AND MAP2K6.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=A;

CC IsoId=Q9Y6R4-1; Sequence=Displayed;

CC Name=B;

CC IsoId=Q9Y6R4-2; Sequence=VSP 004884;

CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN HEART, PLACENTA,

CC SKELETAL MUSCLE AND PANCREAS, AND AT LOWER LEVELS IN OTHER

CC TISSUES.

CC -!- DOMAIN: THE N-TERMINAL NON-CATALYTIC DOMAIN (1-606) CONTAINS AN

CC ESSENTIAL DOMAIN FOR THE DOMINANT-INHIBITORY EFFECT.

CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. MAP

CC kinase kinase subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF002715; AAB6804.1; -.

DR EMBL; D86968; BAA13204.2; ALT_INIT.

DR PIR; T03022; T03022.

DR HSSP; P24941; 1HCL.

DR Gnew; HGNC:6856; MAP3K4.

DR MTM; 602425; -.

DR GO; GO:0004709; F:MAP kinase kinase activity; TAS.

DR GO; GO:0000186; P:activation of MAPKK; TAS.

DR GO; GO:0007254; P:JNK cascade; TAS.

DR GO; GO:0006950; P:response to stress; TAS.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR008271; Ser_thr_pkin_AS.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKC; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW Transferase; Serine/threonine-protein kinase; ATP-binding;

KW Alternative splicing.

FT DOMAIN 4 7 POLY-ALA.

FT DOMAIN 25 38 POLY-PRO.

FT DOMAIN 1190 1201 POLY-ALA.

FT DOMAIN 1342 1600 PROTEIN_KINASE.

FT NP_BIND 1348 1356 ATP (BY SIMILARITY).

FT BINDING 1371 1371 ATP (BY SIMILARITY).

FT ACT_SITE 1462 1462 BY SIMILARITY.

FT VARSPIC 1175 1223 Missing (in isoform B).

FT /FTID=VSP 004884.

FT MUTAGEN 1371 1371 K->R: LOSS OF ACTIVITY.

FT CONFLICT 791 791 I -> R (IN REF. 2).

SQ SEQUENCE 1607 AA; 181550 MW; 24BC4597730B5287 CRC64;

Query Match 7.3%; Score 318.5; DB 1; Length 1607;

Best Local Similarity 23.0%; Pred. No. 1.1e-08;

Matches 116; Conservative 98; Mismatches 161; Indels 129; Gaps 23;

QY 300 PSNDSCELRNKSVQNSHFKEPLVSDKSSSLITDSTLKNKTESLLAKLE----- 353

Db 1206 PSGDSV---LPKSISSAH-----DTRGSSVPENDRL-----ASIAAELQFRSLR 1248

QY 354 ---ETKEYQEPEVE-----SNQKQWQAKRSEBCINQNPAASSNHHQIPELARKVNTQK 405

Db 1249 HSSPTERDEPAYPRGDSGSTRRSWEIR--TLISQSKDTASKLGLPIAIIQKSVR---- 1301

QY 406 HTTFEQPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTVPVKNDFPPACQL 465

Db 1302 --LFEKRYREMRKNIG-----QVCPTPKS--YDNVMHVLKRVT----- 1339

QY 466 STPYGPACFQQQHQILATPLQNLQVLASSANECISVKGRIYSILKQIGSGSSKVQ 525

Db 1340 -----FKWQRG-----NKIGEGYGVVY 1358

QY 526 VLN-EKKQIYAIKYVNLHEDNQTLDSYRNEIAYLNKLQHSKTIIRLYDEITDQYIM 584

Db 1359 CISVDTGELMAMKEIRFPQNDHKTIKETADELKIFEGIKH--PNLVRYFGVELHREEMYI 1416

QY 585 VME-C--GNIDLSWLKKKSIDPWERKSYWKNMLEAVHTIHQGHVHSDLKPAN-FLIV 640

Db 1417 FMEYCDGTLLEVSRLGLEHV----IRLYSKQITIAINVLHEHGVHVDIKGANIFLTS 1472

QY 641 DGMKLIDFGIANQMPDTSVVKD--SQVGTVMYMPPEAIKDNSSSRENGSKSKISPK 698

Db 1473 SGLIKLGFCSVCLKNNAQTMPGEVNSTLGTAAYPEVI-----TRAKGEHGR---A 1524

QY 699 SDVWSLGCILYMTYKTPFPQIINQISKLHAIIDPNHEIEFPDPIEK---DLQDVLKCC 755

Db 1525 ADIWSLGCVVIEMVTGKRPWHEYEHNFQIMYK-VGMGHK---PPIPERLSPEGKDFLSHC 1580

QY 756 LKRDPKQRISIPELLAHPYVQIQ 779

OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98026892; PubMed=9321395;
RA Samejima I., Mackie S., Fantes P.A.;
RT "Multiple modes of activation of the stress-responsive MAP kinase
RT pathway in fission yeast.";
RL EMBO J. 16:6162-6170(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [3]
RP SEQUENCE OF 96-1401 FROM N.A.
RC STRAIN=972;
RX MEDLINE=97282620; PubMed=9136929;
RA Shieh J.C., Wilkinson M., Buck V., Morgan B., Makino K.,
RA Millar J.B.A.;
RT "The Mcs4 response regulator coordinately controls the stress-
RT activated Waki-Wisl-Styl MAP kinase pathway and fission yeast cell
RT cycle.";
RL Genes Dev. 11:1008-1022(1997).
RN [4]
RP SEQUENCE OF 457-543 FROM N.A.
RC STRAIN=972;
RX MEDLINE=97331762; PubMed=9188094;
RA Shiozaki K., Shiozaki M., Russell P.;
RT "Mcs4 mitotic catastrophe suppressor regulates the fission yeast cell
RT cycle through the Waki-Wisl-Spc1 kinase cascade.";
RL Mol. Biol. Cell 8:409-419(1997).
CC -!- FUNCTION: Involved in a signal transduction pathway that is
CC activated in under conditions of heat shock, oxidative stress or
CC limited nutrition. Unlike wisk, it is not activated by changes in
CC the osmolality of the extracellular environment. Activates the
CC wisk MAP kinase kinase by phosphorylation.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. MAP
CC kinase kinase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z98763; CAB11500.1; --
DR EMBL; Y07750; CAA69030.1; ALT_INIT.
DR EMBL; Y11989; CAA72718.1; --
DR EMBL; U81521; AAB39762.1; --
DR PIR; T39225; T39225.
DR HSSP; P24941; ICKP.
DR GeneDB Spombe; SPAC9G1.02; --
DR GO; GO:0007049; P:cell cycle; ISS.
DR GO; GO:0000165; P:MAPKK cascade; ISS.
DR GO; GO:0007165; P:signal transduction; ISS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 1037 1306 PROTEIN_KINASE.
FT NP_BIND 1043 1051 ATP (BY SIMILARITY).
FT BINDING 1066 1066 ATP (BY SIMILARITY).
FT ACT_SITE 1161 1161 BY SIMILARITY.
FT CONFLICT 484 485 RL -> SP (IN REF. 3).
FT CONFLICT 1346 1346 D -> V (IN REF. 1).
SQ SEQUENCE 1401 AA; 160537 MW; 529A989D2B627F3D CRC64;

Query Match 7.0%; Score 304.5; DB 1; Length 1401;
Best Local Similarity 22.1%; Pred. No. 4.5e-08;
Matches 129; Conservative 86; Mismatches 199; Indels 171; Gaps 22;

QY 304 DSCELRLKLS---VQNSHFKEPLVSDKSSSLITDITLKNK----- 343
DB 887 DECDLMDRKTRFWGVGALEFLMLMI--RGNILLIDDDAMFLKIREKVGKSMAPLLTHFDV 944
QY 344 --TESSLLAKLEETKEYQEPEVPESNQKQWAKRKSECINQNPAASSNHWQIPELARKVN 401
DB 945 LGAKSKVAALKQ-----RESTEVSSSPRLTS-FGDVEEALSQ 982
QY 402 TEQKHTTFEQPVFSVSKQSPPISTSKWFDPKSICKTPSSNTLDDYMSCFRTVPVKNDFPP 461
DB 983 LLQKETMLRIDELEIERNTLLE-----RLAIGHVLDD--SVFR---NRDP-- 1023
QY 462 ACOLSTPYGPACFQQQHQHQLATPLQNLQVLASSANECISVKGRISILKQIGSGSS 521
DB 1024 -IKLASSFSNITIRWQGH-----FVRSGMFG 1049
QY 522 KVFQVLN-EKKQIYAIKYVNLEAD--NQTLDSYRNEIAYLNKLOQHSCKIIRLYDYEIT 578
DB 1050 DVTGVNMTGDLAVKEIKLQDSRTFRSTVDQIHNEMTVLERL--NHPNVVYYGVVEVH 1107
QY 579 DQYIYVME-CGNIDLNSWLKSKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLKPNANF 637
DB 1108 REKVYIFMEFCQGGSLADLLAHGRIEDENVLVKVVVQLLEGAYIHSQHILHRDIKPANI 1167
QY 638 LIVD-GMLKLIDFGIANQMOPDTTSVVKDSQV-----GTVMYMPPEAIKDMSSREN 688
DB 1168 LLDHRGMIKYSDFGSALYVSPPTDPEVRYEDIQPELQHLAGTMYMAPEIIL----- 1219
QY 689 GKSKSKISPKSDVWSLGCILYMYTYGTPFQQIINQISKLHAIDPNHEIEFPDIPEKD- 747
DB 1220 GTKKGDFG-AMDIWSLGCVILEMNTGSTPSEMNDNEWAIMHVA----AMHTPSIPQNEK 1274
QY 748 ---LQDVLKCKLRDPKQKRSIPPELLAHYV-----QIQTHPVNQMAKGT----- 790
DB 1275 ISSLARDFIEQCERDEPQRPRAVDILLTHPWITDFRKTIITPPATITKTSLSHTITE 1334

Db	1138	KRQNTXWGTTRMVEVTENHVMVSINKAKN-----SKGEYKEFAWMKG-----EMIGKSFG	1187
Qy	522	KVFQVLN-----EKKQIYAIAKYVNLEEADNQILDSYRNEIAYLNKLQHQHSDKIIRLYD	574
Db	1188	AVYLCNVTTGEMMAVKQVEVPKYSSQNEAILSTVEALRSEVSTLKDLDHLN--IVQYLG	1245
Qy	575	YEITDOYIYMVE-CGNIDLNSWLKKKKSIDPWERKSYWKNMLEAVHTIHQHGIVHSDLK	633
Db	1246	FENKNNIYSLFLEYVAGGSVGLIRMYGRFDEPLIKHLTTQVLKGLAYLHSGILHRDMK	1305
Qy	634	PANFLI-VDGMLKLIDFGIANQMPDTTTSVVKDSQVGTVNYMPPEAIKDMSSSRENGKSK	692
Db	1306	ADNLLDQDGICKISDFGISRKS-K-DIYSNSDMTMRGTVFWMAPMV-----DTK	1354
Qy	693	SKISPKSDVWSLGCILYYMTYKTPFQIINQISKLHAIIDPNHEIEFPDIPEKDL----	748
Db	1355	QYSAKVDIWSLGCIVLEMFAGKRPW-----SNLEVVAAMFKIGKSKSAPPPIPEDTLPLIS	1410
Qy	749	---QDVLKCKLKRDPKQRISEPELLAHPYVQI	777
Db	1411	QIGRNFLDACFEINPEKRPTANELLSHPFSEV	1442

Search completed: September 29, 2004, 16:16:46
Job time : 32 secs